BACHELOR'S THESIS

Degree in Mathematics

Title: Binary Quadratic Forms

Author: Marc Felipe i Alsina

Advisor: Jordi Quer Bosor

Department: Algebraic Number Theory

Academic year: 2017-2018

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Contents

0 Introduction

Integer binary quadratic forms ($ax^2 + bxy + cy^2$) have been studied since the golden era of mathematics.

P. de Fermat and L. Euler, for instance, studied some representations of numbers by particular forms, and characterized them in terms of congruences. Their most well-known theorem in that fashion is the fact that if p is an odd prime, then it can be written as $p = x^2+y^2 \iff p \equiv 1 \mod 4$. Mathematicians of all times have struggled to characterize numbers represented by given quadratic forms, and so quadratic forms have been deeply studied, and for so long. They have numerous appearances and uses in several branches of mathematics, mostly in number theory and algebraic geometry: basic arithmetic, quadratic fields, elliptical curves, among others. An example may be to find units in a quadratic field: one must find all solutions of $N(\varepsilon) = 1$, and N turns out to be a quadratic form when expressing ε in a bssis.

The first mathematician to study quadratic forms as such was J.-L. Lagrange. He was also the first to introduce the concept of equivalent forms, which translated the problem of finding representation by some form to the same problem on a simpler 'reduced' form. A.-M. Legendre noticed that if two numbers are represented by some quadratic forms f and g, the product of those numbers can be represented by a third quadratic form, and so he related multiplication of numbers to a new operation, called composition of forms. However a distinction between equivalence and proper equivalence was still not present at that time, so the operation ended up being multivalued. C. F. Gauß noticed the benefits of dealing with proper equivalence rather than equivalence alone. In his book *Disquisitiones Arithmeticae* [1], which was a great source of inspiration for this degree thesis, he made a great study of quadratic forms, part of which we reproduced here. In particular, his level of abstraction was such, that he managed to give a group structure not to a set of quadratic forms, but to a set of classes of quadratic forms, and all that before the concept of equivalence classes or even the concept of group had been formalized. Later on, the work of Gauß was refined by P. G. L. Dirichlet, who made a new insight on Gauß's composition by connecting it to the theory of ideals, where the same group structure arises. This connection led to what is now known as number field class groups, which has extended beyond the quadratic case.

There is some debate about the definition of an integer quadratic form, as some authors may or may not impose a factor of two in the xy coefficients. Gauß preferred to use $ax^2+2bxy+cy^2$, which is natural since it is the form associated to an integer symmetric bilinear matrix. However, nowadays $ax^2 + bxy + cy^2$ has become the standard, and is the one used in this degree thesis, as well as most modern articles and books about quadratic forms, such as [2],[3],[4] and [5] among others.

More recently, several authors have revisited the works of these great mathematicians and connected them to other areas of mathematics, and some others have rewritten the theory and explained them to the general public, to a greater or lesser extent. We include some of them in the bibliography. In particular, we would like to remark the work of D. A. Cox in his book

Primes of the form $x^2 + ny^2$ [2], who shows the theory of definite forms in great detail and in a very accessible way.

On the other hand, other authors get out of the constraint of using integer coefficients and study binary quadratic forms over a general or concrete ring (or field). Because of the appearance of a factor 2 in the matrix form, the study becomes different when 2 is a zero divisor (or characteristic) in the ring (or field) of coefficients.

In this degree thesis, we present some of the theory of integer binary quadratic forms, namely the classification by discriminant, proper equivalence relation, the composition laws and their connection to the ideal class groups. We will also see how quadratic forms are intrinsically connected to some other areas of mathematics, such as the continued fraction of quadratic numbers, which encodes the transformations between neighbouring half-reduced forms in a cycle, or such as the group of units in a quadratic field, which relates to the study of automorphisms of primitive forms, which in turn represent the solutions of Pell's equation $x^2 - Dy^2 = \pm 4$.

The thesis is meant to be written in a way that any graduate student of mathematics could understand it in its totality.

1 General concepts

First of all, we present what will be the main subject of study in this degree thesis: integer binary quadratic forms.

Definition 1.1. An *integer binary quadratic form* is an homogeneous polynomial of degree 2 on two variables which has integer coefficients: $f(x,y) = ax^2 + bxy + cy^2$, where $a, b, c \in \mathbb{Z}$. It can also be thought to be the result of the following matrix multiplication:

$$f(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + bxy + cy^2; \quad a, b, c \in \mathbb{Z}$$

We may refer to them as *quadratic forms*, or simply *forms*, since we will not study other types of forms.

We adopted the modern definition of integer form. As we stated in the introduction, this definition differs from Gauß's $ax^2 + 2bxy + cy^2$. His definition may seem more restrictive than ours, since it only allows even xy coefficients, but we need to have in mind that the form $ax^2 + bxy + cy^2$ behaves a lot like $2ax^2 + 2bxy + 2cy^2$, so it does not lose much generality.

Definition 1.2. The *content* $\mathrm{Cont}(f)$ of a quadratic form $f(x,y) = ax^2 + bxy + cy^2$ is defined to be the greatest common divisor of a, b and c, which is defined up to the sign.

Definition 1.3. A form is said to be *primitive* if its content is 1.

Definition 1.4. The *discriminant* $D(f)$ of a form $f(x,y) = ax^2 + bxy + cy^2$ is defined as $b^2 - 4ac$.

Gauß [1, §154]defined the discriminant (calling it determinant) to be $b^2 - ac$, which differs from our definition by a factor of 4. Most of the work in our setting can be easily adapted to Gauß's notation by adding a factor in some places, and so, in most occasions it does not matter which definition of integer form is used. However, in some proofs in Gauß's setting it is needed to distinguish between $\gcd(a, 2b, c) = 1$ and $\gcd(a, 2b, c) = 2$, while, in our setting, some other proofs need to be discussed depending on the parity of the xy coefficient, so either way has its pros and cons.

Quadratic forms are intimately connected to bilinear forms, and so to symmetric matrices. This can be seen, for instance, in the definition of quadratic form. The relevant notions can be appreciated from both sides: the discriminant of a form and the determinant of the associated matrix differ only by a factor of -4:

$$b^2 - 4ac = -4 \cdot \det \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

Definition 1.5. A number n is said to be *represented by* a quadratic form f if there exist some integers x_0, y_0 such that $f(x_0, y_0) = n$. A number represented by f is said to be *primitively represented* if, in addition, we have $\gcd(x_0, y_0) = 1$.

It is funny how we use the term 'form' to describe a function of multiple inputs and a single output, mostly used for functions from $\mathbb{R}^n$ to $\mathbb{R}$. The original usage of this word comes from denoting representation using expressions like «primes of the form $4k+1$» or «numbers of the form x^2+y^2», which led Legendre to coin the terms 'linear forms' and 'quadratic forms'.

Example 1.6. The form $2x^2+2xy+3y^2$ is a binary quadratic form of discriminant -20 which is primitive since its content is 1. By letting $x=y=1$, we see that 7 is primitively represented by this form. The number 0 is also represented by this form (in fact, by all forms), but it is not primitively represented (as we will see later).

When talking about quadratic forms, there is one notion that particularly stands out, and that is the notion of equivalence:

Definition 1.7. We say two forms $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ are *equivalent* (and we write $f\sim g$) precisely when there exist some integers $\alpha,\beta,\gamma,\delta\in\mathbb{Z}$ with $\alpha\delta-\beta\gamma=\pm1$ such that $f(\alpha x+\beta y,\gamma x+\delta y)=g(x,y)$. If $\alpha\delta-\beta\gamma=1$, we say that the equivalence is *proper* and *improper* otherwise.

In other words,

$$f\sim g \iff \exists\ \alpha,\beta,\gamma,\delta\in\mathbb{Z}:\ \alpha\delta-\beta\gamma=\pm1 \text{ and}$$

$$\begin{pmatrix}x & y\end{pmatrix}\left[\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{\top}\begin{pmatrix}a & \frac{b}{2}\\ \frac{b}{2} & c\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\right]\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}a' & \frac{b'}{2}\\ \frac{b'}{2} & c'\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} \tag{1.1}$$

This equivalence notion, concerns about invertible linear changes of variables. In the case of the integers, the invertible condition is stated as $\alpha\delta-\beta\gamma=\pm1$, but for forms over a general ring or field, one may impose that $\alpha\delta-\beta\gamma$ is a unit, since this is a necessary and sufficient condition to ensure the change of variables is invertible.

Determining whether two forms are equivalent/properly equivalent or not is rather difficult. We will shed some light on this aspect.

Theorem 1.8. *Both equivalence and proper equivalence are, as the names suggest, equivalence relations.*

Proof. It is easy to see that equivalence and proper equivalence are reflexive and transitive. The symmetric property comes from the fact that $\alpha\delta-\beta\gamma=\pm1$ ensures the change of variables is invertible over the integers:

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} \iff \frac{1}{\alpha\delta-\beta\gamma}\begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}x\\ y\end{pmatrix} \tag{1.2}$$

Therefore, they are indeed equivalence relations. □

In fact, we nowadays talk about an 'equivalence relation' and about 'equivalence classes' on a set because of the usage of the same words in the framework of quadratic forms, which was later exported into similar constructions on other sets. So it is not that we call this relation equivalence because it is an equivalence relation, it is the other way round.

By looking at (1.1), equivalence and proper equivalence classes can be seen as the orbits of the action of $\mathrm{GL}_2(\mathbb{Z})$ and $\mathrm{SL}_2(\mathbb{Z})$, respectively, on the quadratic forms, acting on the right. The action, which can be seen at the level of forms or at the level of symmetric matrices, is described by:

$$\left(f(x,y), \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\right) \mapsto g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$$

$$\left(\begin{pmatrix}a & \frac{b}{2}\\ \frac{b}{2} & c\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\right) \mapsto \begin{pmatrix}a' & \frac{b'}{2}\\ \frac{b'}{2} & c'\end{pmatrix} = \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{\top} \begin{pmatrix}a & \frac{b}{2}\\ \frac{b}{2} & c\end{pmatrix} \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

When considering quadratic forms over a general ring R (or a field $\mathbb{K}$), the groups acting over the set of forms are $\mathrm{GL}_2(R)$ and $\mathrm{SL}_2(R)$ (or $\mathrm{GL}_2(\mathbb{K})$ and $\mathrm{SL}_2(\mathbb{K})$). That is, the group of invertible matrices and the group of matrices with determinant 1.

Anyway, these actions are not faithful: there are two matrices that act identically over all quadratic forms. For example, the action corresponding to the following matrices:

$$\mathrm{Id} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \qquad \text{and} \qquad -\mathrm{Id} = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$$

leave all forms invariant. In that regard, some authors prefer the quotient groups $\mathrm{GL}_2(\mathbb{Z})/\{\mathrm{Id}, -\mathrm{Id}\}$ and $\mathrm{SL}_2(\mathbb{Z})/\{\mathrm{Id}, -\mathrm{Id}\}$ to be the ones who act over the quadratic forms. In the integer case, these groups coincide with $\mathrm{PGL}_2(\mathbb{Z})$ and $\mathrm{PSL}_2(\mathbb{Z})$, respectively.

Since proper equivalence is also a kind of equivalence, equivalence classes consist of a disjoint union of proper equivalence classes. In particular, in the integer case it is the union of at most two of those classes since $\mathrm{GL}_2(\mathbb{Z})/\mathrm{SL}_2(\mathbb{Z}) \cong \mathbb{Z}^\times = \{1, -1\}$ has two elements.

Since equivalence concerns about changing variables, the numbers represented by a form are preserved under equivalence:

Theorem 1.9. *Let $f \sim g$ be two equivalent forms, and let $n \in \mathbb{Z}$. Then, the number of representations and of primitive representations of n by f and g are the same.*

Proof. Since f and g are equivalent, for each representation of n by g, we can find a representation by f since $n = g(x_0, y_0) = f(\alpha x_0 + \beta y_0, \gamma x_0 + \delta y_0) = f(x_0', y_0')$, and vice versa. This correspondence also holds when considering primitive representations, since (1.2) ensures $\gcd(x_0, y_0) \mid \gcd(x_0', y_0')$ and $\gcd(x_0', y_0') \mid \gcd(x_0, y_0)$. □

Apart from the number of representations, there are other invariants under the equivalence relation:

Proposition 1.10. *Equivalent forms have the same discriminant and content.*

Proof. From (1.1), we can find a relation between the coefficients of equivalent quadratic forms $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = a'x^2 + b'xy + c'y^2$:

$$\begin{cases} a' &= a\alpha^2 + b\alpha\gamma + c\gamma^2 = f(\alpha,\gamma) \\ b' &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta \\ c' &= a\beta^2 + b\beta\delta + c\delta^2 = f(\beta,\delta) \end{cases} \tag{1.3}$$

Because of (1.3), we see that $\gcd(a,b,c) \mid \gcd(a',b',c')$ and, by symmetry, it must happen that $\gcd(a',b',c') \mid \gcd(a,b,c)$, so both contents agree. On the other hand, thanks to (1.1), we have that

$$D(g) = (\alpha\delta - \beta\gamma) \cdot D(f) \cdot (\alpha\delta - \beta\gamma) = D(f) \cdot (\pm 1)^2 = D(f)$$

Therefore, the discriminants are the same as well. □

Definition 1.11. If the discriminant is negative, we say that the form is *definite*, if it is the square of a natural number, we say that the form is *degenerate*, and we say it is *indefinite* otherwise. This definition can be formulated at the level of classes as well.

Example 1.12. $2x^2 + 2xy + 3y^2$ (of discriminant -20) is definite, $4x^2 - 10xy + 4y^2$ (of discriminant 36) is degenerate and $x^2 - 2y^2$ (of discriminant 8) is indefinite.

'Degenerate' forms are so-called because they can be factored into linear terms:

Proposition 1.13. *Degenerate forms are precisely those that factor into integer linear forms. They are the only ones that primitively represent* 0*.*

Proof. On the one hand, suppose we have a form like

$$f(x,y) = (\mathrm{A}x + \mathrm{B}y)(\Gamma x + \Delta y) = \mathrm{A}\Gamma x^2 + (\mathrm{A}\Delta + \mathrm{B}\Gamma)xy + \mathrm{B}\Delta y^2, \text{ being } \mathrm{A},\mathrm{B},\Gamma,\Delta \in \mathbb{Z}.$$

Then its discriminant is $(\mathrm{A}\Delta + \mathrm{B}\Gamma)^2 - 4(\mathrm{A}\Gamma)(\mathrm{B}\Delta) = (\mathrm{A}\Delta - \mathrm{B}\Gamma)^2$, so the form is degenerate.

On the other hand, a direct consequence of Gauß's lemma is that if an homogeneous polynomial on two variables has integer coefficients and it is factorable into rational polynomial factors, then the factorization can be modified to feature only integer polynomial factors. Knowing that, if $f(x,y) = ax^2 + bxy + cy^2$ has discriminant $D(f) = b^2 - 4ac = d^2$, then the rational factorization

$$f(x,y) = a\left(x - \frac{-b+d}{2a}y\right)\left(x - \frac{-b-d}{2a}y\right),$$

induces a factorization into integer linear forms. This proves the first part of the proposition.

Since f is degenerate, we can write $f(x,y) = (\mathrm{A}x + \mathrm{B}y)(\Gamma x + \Delta y)$, with Γ and Δ not both zero. Then $f\left(\frac{-\Delta}{\gcd(\Gamma,\Delta)}, \frac{\Gamma}{\gcd(\Gamma,\Delta)}\right) = 0$ is a primitive representation of zero.

Conversely, if $f(x_0, y_0) = 0$ is a primitive representation of 0, in particular x_0 and y_0 are not both zero. Then $f(x, y) = 0$ for all x, y with the same proportion $(\forall\ x : y = x_0 : y_0)$, and therefore f has a linear factor $y_0x - x_0y$ which provides a factorization, a priori, over the rational polynomials. The resulting factorization, if not integer yet, can be modified into an integer factorization of f by linear terms, so f is degenerate. □

The name 'definite' is also not arbitrary. They are called this way because they have a well-defined sign:

Proposition 1.14. *Definite forms only represent either non-negative numbers or non-positive numbers, depending on the sign of a.*

Proof. Indeed, if $f(x, y) = ax^2 + bxy + cy^2$ is definite, then a is non-zero (otherwise $D(f) = b^2 \geq 0$) and $4a \cdot f(x, y) = 4a^2x^2 + 4abxy + 4acy^2 = (2ax + by)^2 - y^2(b^2 - 4ac) \geq 0$. So, the numbers represented by it are always non-negative or non-positive, depending on the sign of a. □

Note that the same argument is valid for c, which necessarily has the same sign of a.

Definition 1.15. We say that a definite form $f(x, y) = ax^2 + bxy + cy^2$ is *positive definite* if $a > 0$ and *negative definite* if $a < 0$.

Finding the number of equivalence classes with a given discriminant is challenging. Since $D = b^2 - 4ac$ is always congruent to 0 or 1 modulo 4, there are no forms for $D \equiv 2, 3 \mod 4$. For the other values, we can see that there is at least one equivalence class:

Definition 1.16. We say that the *principal form* of discriminant D is

$$\begin{cases} x^2 - \frac{D}{4}y^2 & \text{if } D \equiv 0 \mod 4 \\ x^2 + xy - \frac{D-1}{4}y^2 & \text{if } D \equiv 1 \mod 4 \end{cases}$$

We call the proper equivalence class of the principal form the *principal class.*

One of the most remarkable facts about the equivalence classes of quadratic forms is that for each discriminant D there is a finite number of classes. We will prove this result for definite, indefinite and degenerate forms in the following sections.

2 Definite forms

This section broadly follows [2, §2].

Since multiplying a negative definite form by -1 gives a positive definite form, and this transformation respects equivalence and so induces a transformation at the level of classes, we may find the number of equivalence/proper equivalence classes of positive definite forms of discriminant D and multiply the number by 2 to get the total number of classes. Therefore, we will focus on positive definite forms.

Definition 2.1. A positive definite form $f(x,y) = ax^2 + bxy + cy^2$ is called *reduced* if the following inequalities are held:

$$|b| \leq a \leq c, \text{ and } b \geq 0 \text{ whenever } |b| = a \text{ or } a = c$$

Note that the principal form is always reduced. Reduced forms are useful since it will turn out that each positive definite form will be properly equivalent to a unique reduced form. First, let's present an algorithm that, given a positive definite form, will terminate in a properly equivalent reduced form. Therefore, the number of classes cannot be greater than the number of reduced forms. We will start with a positive definite form $f(x,y) = ax^2 + bxy + cy^2$ and we will find a series of properly equivalent forms whose final element is a reduced form.

1. If $|b| > a$, find $k \in \mathbb{Z}$ such that $-a < b + 2ka \leq a$. Make $(x,y) \mapsto (x + ky, y)$, thus getting $ax^2 + (b + 2ka)xy + (ak^2 + bk + c)y^2$. Rename the new coefficients back to a, b and c.

2. If $a > c$, make $(x,y) \mapsto (-y,x)$, thus getting $cx^2 - bxy + ay^2$. Rename the new coefficients back to a, b and c. Start again.

3. If $a = c$ and $b < 0$, make $(x,y) \mapsto (-y,x)$, thus getting $cx^2 - bxy + ay^2$. End.

4. If $b = -a$, make $(x,y) \mapsto (x + y, y)$, thus getting $ax^2 + (b + 2a)xy + (a + b + c)y^2$. End.

5. End.

Example 2.2. We present an example of how the algorithm works. Starting on the form $3x^2 + 8xy + 7y^2$, of discriminant -20, we obtain the following forms:

Step	Form	Rule Applied
0	$3x^2 + 8xy + 7y^2$	1
1	$3x^2 + 2xy + 2y^2$	2
2	$2x^2 - 2xy + 3y^2$	4
3	$2x^2 + 2xy + 3y^2$	

The algorithm stops with $2x^2 + 2xy + 3y^2$ which is reduced.

First, we notice that the algorithm terminated, in our example. It will do so for any form: step 2 reduces the value of a, while step 1 maintains it while lowering $|b|$. Since both a and $|b|$ have to remain non-negative, the process will eventually end.

The resulting form is properly equivalent to the first one since $\alpha\delta - \beta\gamma = 1$ is satisfied by all the transformations in the algorithm. There are three ways to terminate the algorithm. If it ends via the third step, then the resulting form is $ax^2 - bxy + ay^2$ with $0 < -b \leq a$, which is reduced. If it ends via the fourth step instead, we end up with $ax^2 + axy + cy^2$ with $a \leq c$, which is, again, reduced. Finally, if it ends via the fifth step, we have $|b| \leq a \leq c$ with $b \geq 0$ whenever $a = c$ or $|b| = a$, so it is reduced.

Theorem 2.3. *Every positive definite form is properly equivalent to a unique reduced form.*

We have just proved the existence of a reduced form on the orbit of any positive definite form, but we still need to prove uniqueness, so we need to show that two different reduced forms of discriminant D are not properly equivalent. To do so, we will first announce the following lemma:

Lemma 2.4. *If $ax^2 + bxy + cy^2$ is a reduced positive definite form, then*

$$ax^2 + bxy + cy^2 \geq (a - |b| + c)\min\{x^2, y^2\}$$

Proof. Suppose $|x| \geq |y|$ (resp. $|x| \leq |y|$). Then

$$\begin{aligned} ax^2 + bxy + cy^2 &\geq a|x||y| - |b||x||y| + c|y|^2 \geq \\ &\geq (a - |b|)|x||y| + c|y|^2 \geq (a - |b| + c)|y|^2 \end{aligned}$$

$$\begin{aligned} (\text{resp.} \quad ax^2 + bxy + cy^2 &\geq a|x|^2 - |b||x||y| + c|x||y| \geq \\ &\geq a|x|^2 + (c - |b|)|x||y| \geq (a - |b| + c)|x|^2) \end{aligned}$$

□

Now we are able to prove Theorem 2.3:

Proof. (Theorem 2.3) Suppose $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = a'x^2 + b'xy + c'y^2$ are both equivalent and reduced. Suppose $a > a'$. Then, $g(1, 0) = a'$ needs to be primitively represented by f: $f(x_0, y_0) = ax_0^2 + bx_0y_0 + cy_0^2 = a'$. We have $x_0, y_0 \neq 0$ since $a, c > a'$. Then,

$$a' = ax_0^2 + bx_0y + cy_0^2 \geq (a - |b| + c)\min\{x_0^2, y_0^2\} \geq a - |b| + c \geq a > a',$$

which is a contradiction. Therefore $a = a'$ ($a < a'$ is analogously false).

Suppose $c > c'$. Since $g(0, 1) = c'$ is primitively represented by g, it is also primitively represented by f, so now let $f(x_0, y_0) = c'$. But since $a - |b| + c \geq c > c' \geq (a - |b| + c)\min\{x_0^2, y_0^2\}$, we have that either x_0 or y_0 are zero, so $c' \neq c$ must be equal to $a(= a')$, but it is not possible since then the form $g(x, y) = a'x^2 + b'xy + a'y^2$ would primitively represent a' four times:

$(x,y) = (-1,0), (1,0), (0,-1), (0,1)$, but f can only represent it twice if $c \neq a$. Therefore, we have that $c = c'$ ($c < c'$ is analogously false) and we know that f and g can only differ in the xy coefficient.

Since $b^2 = D(f) + 4ac = D(g) + 4a'c' = b'^2$, b and b' can only differ in their sign. Suppose now that $b' \neq b$, so $b' = -b$. Since both f and g are reduced, it must happen that $|b| < a < c$.

Until now we have been using primitive representations as a consequence of f and g being properly equivalent, but in fact the primitively represented numbers are preserved under equivalence alone. We cannot proceed this way since $ax^2 + bxy + cy^2$ and $ax^2 - bxy + cy^2$ are (improperly) equivalent: $(x,y) \mapsto (x,-y)$. However, we may use (1.3) to make sure that all equivalences between f and g must be improper. In this case, we have that:

$$\begin{cases} a &= a\alpha^2 + b\alpha\gamma + c\gamma^2 = f(\alpha,\gamma) \\ -b &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta \\ c &= a\beta^2 + b\beta\delta + c\delta^2 = f(\beta,\delta) \end{cases}$$

Since $a = f(\alpha,\gamma) \geq (a - |b| + c)\min\{\alpha^2,\gamma^2\} \geq a\min\{\alpha^2,\gamma^2\}$, with the last inequality holding with equality only when $\min\{\alpha^2,\gamma^2\} = 0$, we have that $\gamma = 0$ (letting $\alpha = 0$, leads to $a = f(0,\gamma) = c\gamma^2 \geq c > a$ if $\gamma \neq 0$), and therefore $a = f(\alpha,0) = a\alpha^2 \implies \alpha = \pm 1$. Similarly, $c = f(\beta,\delta)$ implies $\beta = 0, \delta = \pm 1$. The fact $\alpha\delta - \beta\gamma = 1$ tells us that α and δ have the same sign. Therefore $b' = -b = 2a\cdot\alpha\cdot 0 + b((\pm 1)^2 + 0\cdot 0) + 2c\cdot 0\cdot\delta = b$, which is a contradiction again.

So, there are not proper equivalences among reduced forms. □

We managed to prove that the number of proper equivalence classes of positive definite forms of a given discriminant is equal to the number of reduced forms of that same discriminant. Moreover, two forms are properly equivalent if they reduce to the same form. As far as equivalence goes, each reduced form $f(x,y) = ax^2 + bxy + cy^2$ is (improperly) equivalent to $g(x,y) = ax^2 - bxy + cy^2$ and, if g is not already reduced, then the algorithm applied to g returns f. So, two forms are equivalent if their reduced forms are equal or differ in the sign of the xy coefficient. To know the number of classes we only need to count how many reduced forms there are. This is not easy, but at least we can say:

Theorem 2.5. *The number of reduced forms of discriminant $D < 0$ is finite, and so is the number of equivalence and proper equivalence classes of that same discriminant.*

Proof. For all reduced forms $ax^2 + bxy + cy^2$, since $-D = 4ac - |b|^2 \geq 4a(a) - a^2 = 3a^2$, we have that $a \leq \sqrt{\frac{-D}{3}}$, so a can have finitely many possible values. $|b| \leq a$ also has finitely many possibilities, and $c = \frac{b^2 - D}{4a}$ is then completely determined, which means there is a finite number of reduced forms, and so of proper equivalence classes. As we said before, the equivalence classes are unions of proper equivalence classes, and they are also finite. □

Therefore, the number of equivalence classes of positive definite forms of discriminant D (without asking for properness) is equal to one half of the sum of the number of proper equiv-

alence classes and the number of reduced forms $ax^2 + bxy + cy^2$ with $b = a$ or $a = c$.

Given $D < 0$, we usually denote by $h(D)$ the number of proper equivalence classes of primitive positive definite forms. We will give a group structure to this set of classes in another section, and $h(D)$ will be the order of the group.

Example 2.6. Using the bijection between proper classes and reduced forms, we can find $h(D)$ for any given D:

D	$h(D)$	D	$h(D)$
-3	1	-28	1
-4	1	-31	3
-7	1	-32	2
-8	1	-35	2
-11	1	-36	2
-12	1	-39	4
-15	2	-40	2
-16	1	-43	1
-19	1	-44	3
-20	2	-47	5
-23	3	-48	2
-24	2	-67	1
-27	1	-163	1

It is known, but it is not easy to prove, that there are no more negative discriminants than the ones presented, whose class number is one.

Before studying degenerate and indefinite forms, it is worth remarking that we can bring back a given indefinite form to its properly equivalent reduced form only making use of two kinds of transformations: the operations $(x, y) \mapsto (x + ky, y)$ and $(x, y) \mapsto (-y, x)$, which are the actions of T^k and S, respectively, where

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

However, this is still not a proof that $\mathrm{SL}_2(\mathbb{Z})$ is generated by these two matrices (which is true) because there are 'automorphisms' that leave a form invariant and they may not be generated by these two matrices, a priori. We will later give a proof that these two matrices generate $\mathrm{SL}_2(\mathbb{Z})$.

3 Indefinite forms

We will now move on towards indefinite forms. Although we could redefine the notion of reduced form using absolute values in order to work for indefinite forms, a similar argument to the one presented would only show the finiteness of the number of classes, but we would not have a way to determine whether two forms are equivalent. In order to solve that, we need to change our notion of reduced:

Definition 3.1. We say that an indefinite form $f(x,y) = ax^2 + bxy + cy^2$ of discriminant D is *half-reduced* if $-b < \sqrt{D} - 2|a| < b < \sqrt{D}$.

Some authors (including Gauß [1, §183], one of the main characters in the history of quadratic forms) use the term 'reduced' to refer to what we call half-reduced forms. Even though there is no risk of confusion since one cannot apply both adjectives to the same form, I think that the term half-reduced is more accurate because it highlights the differences that reduced and half-reduced forms will present.

Half-reduced forms will not be as great as reduced forms are for definite forms: each indefinite form will still be properly equivalent to a half-reduced form, but there will be proper equivalences among some half-reduced forms:

Example 3.2. The form $f(x,y) = x^2 + 8xy - 8y^2$, of discriminant 96 is half-reduced since $-8 < \sqrt{96} - 2 < 8 < \sqrt{96}$. The same thing happens with the form $g(x,y) = -8x^2 + 8xy + y^2$ of the same discriminant, since $-8 < \sqrt{96} - 16 < 8 < \sqrt{96}$. However, these two forms are properly equivalent since $f(-y, x+y) = g(x,y)$.

Another difference with reduced forms, is that the principal form is not half-reduced, but it is easy to find a half-reduced form in the principal class:

Theorem 3.3. *Let $D > 0$ and let k be the greatest integer such that D and k have the same parity and that $k < \sqrt{D}$. Let $l = \frac{D-k^2}{4}$. Then the form $p(x,y) = x^2 + kxy - ly^2$ is half-reduced and it is properly equivalent to the principal form of discriminant D.*

Proof. First of all, note that the condition on the parity forces l to be an integer. By construction, this form has discriminant D. Therefore, by applying the transformation $(x,y) \mapsto$ $\mapsto (x+y, y)$ to the principal form, we will eventually get to $p(x,y)$, since this transformation leaves the first coefficient fixed, increases by two the middle coefficient and modifies the third one in order to preserve the discriminant.

Therefore, we only need to show it is half-reduced, that is: $-k < \sqrt{D} - 2 < k < \sqrt{D}$. The last two inequalities hold because of the election of k. We have that $\sqrt{D} - 2$ is always positive, since D must be positive, congruent to 0 or 1 modulo 4 and not a square, so D is at least 5. Then, k is also positive and $-k$ negative therefore, so the first inequality is also fulfilled. □

As it happens with reduced forms, there is a finite number of half-reduced forms:

Proposition 3.4. *There is a finite number of half-reduced forms of discriminant $D > 0$.*

Proof. Suppose $f(x,y) = ax^2 + bxy + cy^2$ has discriminant D and is half-reduced. From the definition of half-reduced, we have $0 < b < \sqrt{D}$. Then, $-\sqrt{D} < -b < \sqrt{D} - 2|a|$ implies $|a| < \sqrt{D}$ and c is completely determined by a, b and D. Therefore, there can only exist a finite number of half-reduced forms. □

Our objective now is, as we did for the definite case, present an algorithm that will transform a given form into a properly equivalent half-reduced one. To do so, we need to introduce another concept:

Definition 3.5. We say that a form $g(x,y) = a'x^2 + b'xy + c'y^2$ is a *successor* of the form $f(x,y) = ax^2 + bxy + cy^2$ if the following three conditions are met:

- $D(f) = D(g)$
- $a' = c$
- $b + b' \equiv 0 \mod 2c$

In that case, we also say that f is a *predecessor* of g and that f and g are *neighbours*.

Lemma 3.6.
If $g(x,y) = a'x^2 + b'xy + c'y^2$ is a successor of $f(x,y) = ax^2 + bxy + cy^2$, then there exists some $l \in \mathbb{Z}$ such that $g(x,y) = cx^2 + (-b + 2lc)xy + (l^2c - lb + a)y^2$.
If $f(x,y) = ax^2 + bxy + cy^2$ is a predecessor of $g(x,y) = a'x^2 + b'xy + c'y^2$, then there exists some $l \in \mathbb{Z}$ such that $f(x,y) = (l^2a' - lb' + c')x^2 + (-b' + 2la')xy + a'y^2$.

Proof. For the first part, we have that $a' = c$. Let $l = \frac{b+b'}{2c}$, which is an integer since f and g are neighbours. Then, the result follows from $b' = -b + 2lc$ and

$$c' = \frac{b'^2 - D(g)}{4a'} = \frac{b^2 - D(f) - 4lbc + 4l^2c^2}{4c} = l^2c - lb + a$$

The second part is analogous. □

Proposition 3.7. *Each form is properly equivalent to all its neighbours.*

Proof. Suppose that $f(x,y) = ax^2 + bxy + cy^2$ has a successor g, which we can write as $g(x,y) = cx^2 + (-b + 2lc)xy + (l^2c - lb + a)y^2$. The change $(x,y) \mapsto (-y, x + ly)$ makes f and g properly equivalent since

$$a(-y)^2 + b(-y)(x + ly) + c(x + ly)^2 = cx^2 + (-b + 2lc)xy + (l^2c - lb + a)y^2$$

□

It is worth noting that any transformation of the form $(x,y) \mapsto (-y, x + ly)$ will map $ax^2 + bxy + cy^2$ to a successor. So, the set of successors of $ax^2 + bxy + cy^2$ can be parameterized by the set of integers.

We will now present the aforesaid algorithm, which, given an indefinite form $f(x,y) = = ax^2+bxy+cy^2$ of discriminant D, will present a series of forms equivalent to it, but instead of ending and outputting a half-reduced form, it will loop over a cycle of half-reduced forms:

1. Given $f(x,y) = ax^2+bxy+cy^2$,
   - If $\sqrt{D} < |c|$, find the successor $a'x^2+b'xy+c'y^2$ such that $-|c| < b' \leq |c|$
   - If $|c| < \sqrt{D}$, find the successor $a'x^2+b'xy+c'y^2$ such that $\sqrt{D}-2|c| < b' < \sqrt{D}$
2. Rename the new coefficients back to a, b and c. Start again.

The algorithm is well-defined because, since b' is determined modulo $2c$ and there is only one number congruent to it in a $2|c|$-long interval.

Example 3.8. Starting on the form $2x^2+11xy+8y^2$, of discriminant 57, we obtain the following forms using the algorithm:

Step	Form	Rule Applied
0	$2x^2+11xy+8y^2$	1
1	$8x^2+5xy-y^2$	2
2	$-x^2+7xy+2y^2$	2
3	$2x^2+5xy-4y^2$	2
4	$-4x^2+3xy+3y^2$	2
5	$3x^2+3xy-4y^2$	2
6	$-4x^2+5xy+2y^2$	2
7	$2x^2+7xy-y^2$	2
8	$-x^2+7xy+2y^2$	2
9	$2x^2+5xy-4y^2$	2
$\vdots$	$\vdots$	$\vdots$

In this case, the algorithm passes by and then cycles over a series of forms. The ones on the cycle happen to be all half-reduced.

Let's prove a couple of lemmas, concerning the production of half-reduced forms by the algorithm:

Lemma 3.9. *If an indefinite form $f(x,y) = ax^2+bxy+cy^2$ of discriminant D satisfies that $|c| < \frac{\sqrt{D}}{2}$, then one step of the algorithm yields a half-reduced form.*

Proof. Since $|c| < \frac{\sqrt{D}}{2} < \sqrt{D}$, the algorithm will present a form $a'x^2+b'xy+c'y^2$ such that $0 < \sqrt{D}-2|a'| < b' < \sqrt{D}$. The last two inequalities hold because of how the algorithm works

and the first inequality comes from the fact that $\sqrt{D} - 2|a'| = \sqrt{D} - 2|c| > 0$ by hypothesis. The only extra condition we need to check is that $-b' < \sqrt{D} - 2|a'|$, which is true since the right hand side is positive and the left hand side is negative. Therefore, $a'x^2 + b'xy + c'y^2$ is half-reduced. □

Lemma 3.10. *If an indefinite form $f(x, y) = ax^2 + bxy + cy^2$ of discriminant D satisfies $|c| < \sqrt{D}$, then two steps of the algorithm yield a half-reduced form.*

Proof. Let $g(x, y) = a'x^2 + b'xy + c'y^2$ the neighbour of f provided by the algorithm. Let $h(x, y) = a''x^2 + b''xy + c''y^2$ be what we get after applying one step of the algorithm to g. We need to show that h is half-reduced.

Since $|c| < \sqrt{D}$ the algorithm will first be applied under the second condition and we have $\sqrt{D} - 2|a'| < b' < \sqrt{D}$. So, it happens that $0 < \sqrt{D} - b' < 2|a'|$. On the other hand, $-\sqrt{D} < \sqrt{D} - 2|c| = \sqrt{D} - 2|a'| < b'$, so $|b'| < \sqrt{D}$ and $0 < D - b'^2 = -4a'c' = 4|a'||c'|$. This implies that

$$\frac{\sqrt{D} + b'}{2|c'|} = \frac{2|a'|}{\sqrt{D} - b'} > 1,$$

which means $-b' + 2|c'| < \sqrt{D}$. Since $b' < \sqrt{D}$, we have that $|c'| < \sqrt{D}$, so the algorithm will be executed again under the second condition. If $|c'| < \frac{\sqrt{D}}{2}$, we apply Lemma 3.9 to g and we are done. Suppose now that $|c'| > \frac{\sqrt{D}}{2}$. Then

$$-(-b' + 2|c'|) = b' - 2|c'| < \sqrt{D} - 2|c'| < 0 < -b' + \sqrt{D} < -b' + 2|c'| < \sqrt{D},$$

which means that the algorithm will pick $b'' = -b' + 2|c'|$ and the resulting form h will satisfy

$$-b'' < \sqrt{D} - 2|a''| < b'' < \sqrt{D}$$

and will therefore be half-reduced. □

Theorem 3.11. *The algorithm satisfies the following:*

1. *At each step, it produces a form which is properly equivalent to the initial form.*
2. *It will eventually hit a half-reduced form.*
3. *Once we get a half-reduced form, all subsequent forms are also half-reduced.*
4. *The algorithm will eventually hit a form that was produced before and, therefore, will cycle.*

Proof.

1. Since neighbours are properly equivalent, this is obviously true.

2. Suppose that at some point, we get to a form $ax^2+bxy+cy^2$ with $|c|<\sqrt{D}$. Then, Lemma 3.10 implies that we get a half-reduced form and we are done. So now we need to show that this condition will be met eventually. Suppose that $|c|>\sqrt{D}$ and we apply the algorithm once to get $a'x^2+b'xy+c'y^2$. We know that $|b'|\leq|c|$, since the algorithm ran under the first condition, but it may happen that $|b'|<\sqrt{D}$ or $\sqrt{D}<|b'|$. If it is the first case, then $D-b'^2>0$ and $|c'|=\frac{D-b'^2}{4|c|}\leq\frac{D}{4|c|}<\frac{|c|}{4}$. If it is the second case, then also $|c'|=\frac{b'^2-D}{4|c|}\leq\frac{c^2-D}{4|c|}<\frac{c^2}{4|c|}=\frac{|c|}{4}$.

   Therefore, the value of the y^2 coefficient cannot stay over $\sqrt{D}$ indefinitely and we will eventually reach a half-reduced form, making the statement true.

3. Suppose $g(x,y)=a'x^2+b'xy+c'y^2$ is half-reduced. Then, $-b'<\sqrt{D}-2|a'|<b'<\sqrt{D}$ and so $|a'|<\sqrt{D}$. Let $f(x,y)=ax^2+bxy+cy^2$ be a predecessor of g. Then $|c|=|a'|<$ $<\sqrt{D}$ and applying one step of the algorithm to f yields g. Lemma 3.10 applied to f tells us that applying another step to g gives a half-reduced form.

4. The algorithm will eventually hit a half-reduced form, and will produce half-reduced forms from that point on. Since there is a finite number of half-reduced forms of any particular discriminant, a cycle of half-reduced forms will be produced.

□

Corollary 3.12. *There is a finite number of equivalence and proper equivalence classes of indefinite forms of discriminant D.*

Proof. We have proved that for each indefinite form f, there is a half-reduced form (properly) equivalent to it, so the number of equivalence/proper equivalence classes cannot be greater than the number of half-reduced forms, which is finite. □

The algorithm will not only be useful to prove the finiteness of the number of classes, but also to determine whether two forms are properly equivalent, thanks to the following theorem:

Theorem 3.13. *If $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ are both properly equivalent and half-reduced, then applying the algorithm to f will eventually get to g.*

This theorem is hard to prove, and we will have to show some results first:

Proposition 3.14. *If $f(x,y)=ax^2+bxy+cy^2$ is half-reduced, then $h(x,y)=cx^2+bxy+ay^2$ is also half-reduced.*

Proof. Let $D=D(f)=D(h)$. We know that $-b<\sqrt{D}-2|a|<b<\sqrt{D}$, which is equivalent to $\sqrt{D}+b>2|a|>\sqrt{D}-b>0$. Since $-2a\cdot 2c=D-b^2=\left(\sqrt{D}+b\right)\left(\sqrt{D}-b\right)>0$, it happens that $2|a|\cdot 2|c|=\left(\sqrt{D}+b\right)\left(\sqrt{D}-b\right)$, but as $2|a|$ lies in the middle of the two factors, $2|c|$ must also lie there. Therefore, the following inequalities are fulfilled:

$$\sqrt{D}+b>2|c|>\sqrt{D}-b>0 \implies -b<\sqrt{D}-2|c|<b<\sqrt{D}$$

and h is half-reduced. $\square$

Proposition 3.15. *Each half-reduced form has a unique half-reduced predecessor and a unique half-reduced successor.*

Proof. Let $f(x,y) = ax^2 + bxy + cy^2$ be half-reduced. Let g be a successor of f, which we can write as $g(x,y) = a'x^2 + b'xy + c'y^2 = cx^2 + (-b + 2lc)xy + (l^2c - lb + a)y^2$. Then, there exists a unique l that makes b' lie on the range $\sqrt{D} - 2|a'| < b' < \sqrt{D}$, where it must lie in order for g to be half-reduced. So, f has at most one half-reduced successor. Similarly, f only has at most one half-reduced predecessor. The algorithm shows that f has at least one half-reduced successor. The algorithm applied to $h(x,y) = cx^2 + bxy + ay^2$ (which is half-reduced) yields some other half-reduced form $ax^2 + Bxy + Cy^2$. Then the form $Cx^2 + Bxy + ay^2$ is a half-reduced predecessor of f. Therefore, f has a unique half-reduced predecessor and a unique half-reduced successor. $\square$

This last proposition ensures that the cycle of half-reduced forms that we get when using the algorithm starts on the first half-reduced form encountered. Every half-reduced form is part of a cycle.

We are making our way to prove Theorem 3.13, but first we need to introduce some new parameters in regard to a cycle of half-reduced forms. The rest of the proof mainly follows Gauß's proof in [1, §§188-193]:

Given a half-reduced form f_0 of discriminant D, define $(f_i)_{i\in\mathbb{Z}}$ such that f_{i+1} and f_{i-1} are, respectively, the half-reduced successor and predecessor of f_i for all $i \in \mathbb{Z}$. Also, for each integer i, let l_i be the integer such that $f_{i-1}(-y, x + l_i y) = f_i(x,y)$, let $\alpha_0 = 1, \beta_0 = 0, \gamma_0 = 0, \delta_0 = 1$ and let $(\alpha_i)_{i\in\mathbb{Z}}, (\beta_i)_{i\in\mathbb{Z}}, (\gamma_i)_{i\in\mathbb{Z}}, (\delta_i)_{i\in\mathbb{Z}}$ be defined such that

$$\begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_{i+1} \end{pmatrix} = \begin{pmatrix} \alpha_{i+1} & \beta_{i+1} \\ \gamma_{i+1} & \delta_{i+1} \end{pmatrix} \qquad \forall i \in \mathbb{Z} \iff$$

$$\iff \begin{cases} \alpha_{i+1} &= \beta_i \\ \beta_{i+1} &= -\alpha_i + \beta_i l_{i+1} \\ \gamma_{i+1} &= \delta_i \\ \delta_{i+1} &= -\gamma_i + \delta_i l_{i+1} \end{cases} \qquad \forall i \in \mathbb{Z} \tag{3.1}$$

Lastly, let's define $(a_i)_{i\in\mathbb{Z}}$, $(b_i)_{i\in\mathbb{Z}}$ and $(c_i)_{i\in\mathbb{Z}}$ such that $f_i(x,y) = a_i x^2 + b_i xy + c_i y^2$ for all $i \in \mathbb{Z}$.

These definitions ensure that $f_0(\alpha_i x + \beta_i y, \gamma_i x + \delta_i y) = f_i(x,y)$ and also that $\alpha_i\delta_i - \beta_i\gamma_i = 1$ $\forall i \in \mathbb{Z}$. It is clear that the sequence of forms $(f_i)_{i\in\mathbb{Z}}$ is periodic, and so are $(l_i)_{i\in\mathbb{Z}}$, $(a_i)_{i\in\mathbb{Z}}$, $(b_i)_{i\in\mathbb{Z}}$ and $(c_i)_{i\in\mathbb{Z}}$, but this does not need to be true for $(\alpha_i)_{i\in\mathbb{Z}}, (\beta_i)_{i\in\mathbb{Z}}, (\gamma_i)_{i\in\mathbb{Z}}$ and $(\delta_i)_{i\in\mathbb{Z}}$. Also, we can rewrite Equation (3.1) as

$$\begin{cases} \alpha_{i+2} + \alpha_i &= \alpha_{i+1} l_{i+1} \\ \beta_{i+2} + \beta_i &= \beta_{i+1} l_{i+2} \\ \gamma_{i+2} + \gamma_i &= \gamma_{i+1} l_{i+1} \\ \delta_{i+2} + \delta_i &= \delta_{i+1} l_{i+2} \end{cases} \qquad \forall i \in \mathbb{Z} \tag{3.2}$$

Example 3.16. Before going further, let us reuse Example 3.8 to illustrate these new concepts:

i	l_i	f_i	α_i	β_i	γ_i	δ_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
-7	3	$2x^2+7xy-y^2$	2117	-291	291	-40
-6	-7	$-x^2+7xy+2y^2$	-291	-80	-40	-11
-5	3	$2x^2+5xy-4y^2$	-80	51	-11	7
-4	-1	$-4x^2+3xy+3y^2$	51	29	7	4
-3	1	$3x^2+3xy-4y^2$	29	-22	4	-3
-2	-1	$-4x^2+5xy+2y^2$	-22	-7	-3	-1
-1	3	$2x^2+7xy-y^2$	-7	1	-1	0
0	-7	$-x^2+7xy+2y^2$	1	0	0	1
1	3	$2x^2+5xy-4y^2$	0	-1	1	3
2	-1	$-4x^2+3xy+3y^2$	-1	1	3	-4
3	1	$3x^2+3xy-4y^2$	1	2	-4	-7
4	-1	$-4x^2+5xy+2y^2$	2	-3	-7	11
5	3	$2x^2+7xy-y^2$	-3	-11	11	40
6	-7	$-x^2+7xy+2y^2$	-11	80	40	-291
7	3	$2x^2+5xy-4y^2$	80	251	-291	-963
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

We now state and prove a technical lemma about the sign and magnitude of the sequences we have defined. It is recommended to read its proof while looking at Example 3.16.

Lemma 3.17. *The sequences* $(|\beta_i|)_{i\geq 1}, (|\gamma_i|)_{i\geq 1}, (|\beta_{-i}|)_{i\geq 1}$ *and* $(|\gamma_{-i}|)_{i\geq 1}$ *are increasing and not bounded. Also,* $a_0\alpha_i\gamma_i$ *and* $a_0\beta_i\delta_i$ *are non-negative quantities for* $i \geq 1$ *and they are non-positive for* $i \leq -1$.

Proof. We see that a_i and c_i alternate signs: being half-reduced implies $0 < b_i < \sqrt{D}$, which means $4a_ic_i = b_i^2 - D < 0$ and we have $a_{i+1} = c_i$. Moreover, since b_i is always positive, we have that $-b_i + 2l_{i+1}c_i = b_{i+1}$ implies that $2l_{i+1}c_i$ is always positive. Therefore, l_{i+1} also alternates signs, and has the same sign as $c_i = a_{i+1}$.

Note that this also proves that the length of the cycle is always even and that $l_i \neq 0 \quad \forall i \in \mathbb{Z}$.

We will now make a proof by induction, so let's focus on what are the values of β_i and γ_i for low $|i|$'s. By using the formulas in (3.2), we have:

$$\begin{aligned}
\beta_{-3} &= l_0l_{-1} - 1, & \gamma_{-3} &= -l_{-1}l_{-2} + 1\\
\beta_{-2} &= l_0, & \gamma_{-2} &= -l_{-1}\\
\beta_{-1} &= 1, & \gamma_{-1} &= -1\\
\beta_0 &= 0, & \gamma_0 &= 0\\
\beta_1 &= -1, & \gamma_1 &= 1\\
\beta_2 &= -l_2, & \gamma_2 &= l_1\\
\beta_3 &= -l_2l_3 + 1, & \gamma_3 &= l_1l_2 - 1
\end{aligned}$$

We want to prove that, for all $i \geq 1$, we have that β_i and $\beta_{i+1}l_{i+2}$ have opposite signs and also that $|\beta_{i+1}| \geq |\beta_i|$. This statement is true for $i = 1$, as can be seen above. Suppose now that β_i and $\beta_{i+1}l_{i+2}$ have opposite signs and that $|\beta_{i+1}| \geq |\beta_i|$. Then $\beta_{i+2} = -\beta_i + \beta_{i+1}l_{i+2}$ will have a sign opposite to that of β_i, which means that $\beta_{i+2}l_{i+3}$ will have the same sign as $\beta_i l_{i+2}$, whose sign is opposite to that of $(\beta_{i+1}l_{i+2})l_{i+2} = \beta_{i+1}l_{i+2}^2$ and β_{i+1}. Also, $|\beta_{i+2}| = |\beta_i| + |\beta_{i+1}l_{i+2}| \geq 0 + |\beta_{i+1}||l_{i+2}| \geq |\beta_{i+1}|$, which completes the induction step. Therefore, the statement has been proved by induction. Note that the inequality $|\beta_{i+2}| \geq |\beta_{i+1}|$ is strict when $|\beta_i| > 0$.

From this, we can say that β_i and β_{i+2} have opposite signs and that $|\beta_i|$ is increasing for $i \geq 1$ and not bounded (since $|\beta_i| \geq |\beta_1| = 1 > 0$).

Similarly, one can make a similar induction to derive the analogous results for the sequences γ_i, β_{-i} and γ_{-i}.

The only thing that we need to prove is the sign of $a_0\alpha_i\gamma_i$ and $a_0\beta_i\delta_i$. In fact, they are the same sequence shifted, since $\alpha_{i+1} = \beta_i$ and $\gamma_{i+1} = \delta_i$, so we may focus only on the second one: $(m_i)_{i\in\mathbb{Z}}$ where $m_i := a_0\beta_i\gamma_{i+1}$.

For $i = 0, -1$; $m_i = 0$. For $i = -3, -2$; m_i equals $-[-a_0l_{-1}][-(l_0l_{-1} - 1)]$ and $-[a_0l_0]$, respectively. For $i = 1, 2$; m_i equals $[-a_0l_1]$ and $[a_0l_2][-(l_1l_2 - 1)]$, respectively, which fulfill the sign requirements since each expression in square brackets is positive. Then, since $\beta_i\gamma_{i+1}$ and $\beta_{i+2}\gamma_{i+3}$ have the same sign and also do $\beta_{-i-2}\gamma_{-i-1}$ and $\beta_{-i}\gamma_{-i+1}$ for $i \geq 1$, the sign of m_{i+2} will be the same as the one of m_i for $i \geq 1$ and the sign of m_{i-2} will be the same as the one of m_i for $i \leq -2$, which finishes the proof. □

Corollary 3.18. *$\alpha_i \neq 0$, $\beta_i \neq 0$, $\gamma_i \neq 0$, $\delta_i \neq 0$ for all $i \in \mathbb{Z}$, except for α_1, β_0, γ_0 and δ_{-1}.*

Proof. This comes from the fact that α_i and δ_i are shifts of the sequences β_i and γ_i, that $|\beta_1| = |\beta_{-1}| = |\gamma_1| = |\gamma_{-1}| = 1$ and that $|\beta_i|, |\gamma_i|, |\beta_{-i}|$ and $|\gamma_{-i}|$ are increasing for $i \geq 1$. □

Lemma 3.19. *Let $f(x, y) = ax^2 + bxy + cy^2$ be a half-reduced indefinite form of discriminant D and let $g(x, y) = a'x^2 + b'xy + c'y^2 = f(\alpha x + \beta y, \gamma x + \delta y)$ be equivalent to it and also half-reduced. If $\alpha, \beta, \gamma, \delta \neq 0$ and either $\frac{\alpha}{\gamma}$ or $\frac{\beta}{\delta}$ has the same sign as a, the value $\frac{-b+\sqrt{D}}{2a}$ lies in between. If one of them has opposite sign to that of a, the value $\frac{-b-\sqrt{D}}{2a}$ lies in the middle instead.*

Proof. First, we note that if one of the fractions $\frac{\alpha}{\gamma}$, $\frac{\beta}{\delta}$ is positive, the other one is non-negative, and if one of them is negative, the other one is non-positive. This is because their difference, in absolute value, is

$$\frac{1}{|\gamma||\delta|} \leq \min\left\{\left|\frac{\alpha}{\gamma}\right|, \left|\frac{\beta}{\delta}\right|\right\}$$

Using the relation about the coefficients of equivalent forms (1.3), we know that $a' = = f(\alpha, \gamma) = a\alpha^2 + b\alpha\gamma + c\gamma^2$ and $c' = f(\beta, \delta) = a\beta^2 + b\beta\delta + c\delta^2$. If the first hypothesis of the statement happens, we divide by γ^2 and δ^2, respectively. Solving, we obtain that

$$\frac{\alpha}{\gamma} = \frac{-b + \sqrt{b^2 - 4a\left(c - \frac{a'}{\gamma^2}\right)}}{2a} = \frac{-b + \sqrt{D + 4\frac{aa'}{\gamma^2}}}{2a}$$

$$\frac{\beta}{\delta} = \frac{-b + \sqrt{b^2 - 4a\left(c - \frac{c'}{\delta^2}\right)}}{2a} = \frac{-b + \sqrt{D + 4\frac{ac'}{\delta^2}}}{2a}$$

where the sign of the root has been determined since $-b$ is negative but $2a\frac{\alpha}{\gamma}$ and $2a\frac{\beta}{\delta}$ are non-negative.

Similarly, if the second scenario takes place, we divide by α^2 and β^2, respectively. Solving, we obtain that

$$\frac{\gamma}{\alpha} = \frac{-b + \sqrt{b^2 - 4c\left(a - \frac{a'}{\alpha^2}\right)}}{2c} = \frac{-b + \sqrt{D + 4\frac{ca'}{\alpha^2}}}{2c}$$

$$\frac{\delta}{\beta} = \frac{-b + \sqrt{b^2 - 4c\left(a - \frac{c'}{\beta_i}\right)}}{2c} = \frac{-b + \sqrt{D + 4\frac{cc'}{\beta^2}}}{2c}$$

where the sign of the root has been determined since $-b$ is negative but $2c\frac{\gamma}{\alpha}$ and $2c\frac{\delta}{\beta}$ are non-negative since $ac < 0$ because f is half-reduced.

The first result follows from the fact that $\frac{aa'}{\gamma^2}$ and $\frac{ac'}{\delta^2}$ have opposite signs (g is half-reduced), so the quantity $\frac{-b+\sqrt{D}}{2a}$ lies in between the fractions. Similarly, the second case leads to $\frac{-b+\sqrt{D}}{2c}$ lying in the middle of $\frac{\gamma}{\alpha}$, $\frac{\delta}{\beta}$. Therefore, $\frac{2c}{-b+\sqrt{D}} = \frac{-b-\sqrt{D}}{2a}$ will lie in the middle of $\frac{\alpha}{\gamma}$ and $\frac{\beta}{\delta}$. □

Proposition 3.20. *Let D be the discriminant of f_0 (and all the other f_i's). The sequences $\left(\frac{\alpha_i}{\gamma_i}\right)_{i\geq 1}$ and $\left(\frac{\beta_i}{\delta_i}\right)_{i\geq 1}$ tend to $\frac{-b_0+\sqrt{D}}{2a_0}$. The sequences $\left(\frac{\alpha_{-i}}{\gamma_{-i}}\right)_{i\geq 1}$ and $\left(\frac{\beta_{-i}}{\delta_{-i}}\right)_{i\geq 2}$ tend to $\frac{-b_0-\sqrt{D}}{2a_0}$ instead. Each of the successions has every other term over the limit and every other under it. Moreover, the odd and even subsequences are monotonic.*

Proof. We will prove the result for the sequence $\left(\frac{\alpha_i}{\gamma_i}\right)_{i\geq 1}$. It automatically will be proved for the sequence $\left(\frac{\beta_i}{\delta_i}\right)_{i\geq 1}$, which is just a shift of the other sequence. The analogous results concerning the other sequences can be proved using the same arguments.

We have that $\frac{\alpha_i}{\gamma_i}$ and $\frac{\beta_i}{\delta_i} = \frac{\alpha_{i+1}}{\gamma_{i+1}}$ have the same sign as a_0 for $i \geq 1$, since $a_0\alpha_i\gamma_i$ is non-negative for $i \geq 1$. So, $\frac{-b_0+\sqrt{D}}{2a_0}$ lies between the two fractions. Since $\left|\frac{\alpha_i}{\gamma_i} - \frac{\alpha_{i+1}}{\gamma_{i+1}}\right| = \left|\frac{\alpha_i\delta_i - \beta_i\gamma_i}{\gamma_i\delta_i}\right| = \frac{1}{|\gamma_i||\gamma_{i+1}|}$ tends to zero, this intermediate value must be its limit.

With respect to the second statement, note that what determines if $\frac{\alpha_i}{\gamma_i}$ over- or undershoots its limit is whether $4\frac{a_0 a_i}{\gamma_i^2}$ has the same sign as $2a_0$ or not. This condition only depends on the sign of a_i and so, of the parity of i, as we wanted to prove. The fact that $\alpha_i\gamma_{i+1} - \alpha_{i+1}\gamma_i = \alpha_i\delta_i - \beta_i\gamma_i = 1$ implies that every fraction in the interval with endpoints $\frac{\alpha_i}{\gamma_i}$ and $\frac{\alpha_{i+1}}{\gamma_{i+1}}$ (which contains $\frac{-b_0+\sqrt{D}}{2a_0}$) can only contain fractions with greater denominator that $|\gamma_i|$ and $|\gamma_{i+1}|$. Therefore, the fractions $\frac{\alpha_j}{\gamma_j}$ with $j < i$ lie outside it. That is, a larger i implies the fraction is closer to the limit, ensuring the monotonicity of the subsequences. □

The situation is, therefore, the following:

$$\cdots < \frac{\alpha_i}{\gamma_i} < \frac{\alpha_{i+2}}{\gamma_{i+2}} < \frac{\alpha_{i+4}}{\gamma_{i+4}} < \cdots < \frac{-b_0+\sqrt{D}}{2a_0} < \cdots < \frac{\alpha_{i+3}}{\gamma_{i+3}} < \frac{\alpha_{i+1}}{\gamma_{i+1}} < \cdots$$

$$\cdots < \frac{\alpha_{-i}}{\gamma_{-i}} < \frac{\alpha_{-i-2}}{\gamma_{-i-2}} < \frac{\alpha_{-i-4}}{\gamma_{-i-4}} < \cdots < \frac{-b_0-\sqrt{D}}{2a_0} < \cdots < \frac{\alpha_{-i-3}}{\gamma_{-i-3}} < \frac{\alpha_{-i-1}}{\gamma_{-i-1}} < \cdots$$

Now, we are in conditions to prove Theorem 3.13:

Proof. (Theorem 3.13) Let $f_0(x,y) = f(x,y) = ax^2 + bxy + cy^2$ and define accordingly all parameters used in the former lemmas. We need to show that $g(x,y) = a'x^2 + b'xy + c'y^2$ is equal to one of the f_i's. Since f and g are properly equivalent, there exist some integers α, β, γ and δ with $\alpha\delta - \beta\gamma = 1$ such that $f(\alpha x + \beta y, \gamma x + \delta y) = g(x,y)$.

Suppose $\alpha = 0$. We have that $\beta\gamma = -1$, which implies $\pm\beta = -1$, $\pm\gamma = 1$. Since $(x,y) \mapsto (\pm\alpha x \pm \beta y, \pm\gamma x \pm \delta y) = (-y, x \pm \delta y)$ transforms f into successor (because of the particular form of the transformation) but also to a half-reduced form g, it must happen that g is the half-reduced successor of f. That is, $g = f_1$.

Suppose $\beta = 0$. Then, $\alpha = \pm 1$, $\delta = \pm 1$. We have that $c' = f(\beta,\delta) = f(0,\pm 1) = c$, and $b' = 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta = b + 2c\gamma\delta$, so b and b' are congruent modulo $2c$. These conditions imply that both f and g are half-reduced predecessors of f_1, so $g = f = f_0$.

Similar arguments hold for the cases $\gamma = 0$, $\delta = 0$. We will now suppose that $\alpha, \beta, \gamma, \delta \neq 0$.

The idea under the rest of the proof is to show that α, β, γ and δ equal $\alpha_i, \beta_i, \gamma_i$ and δ_i, respectively, for some i. To prove that, we first show that $\frac{\alpha}{\gamma} = \frac{\alpha_i}{\gamma_i}$ for some i, and we do that by *reductio ad absurdum*. We will constantly make use of the fact that if $\mathrm{A}\Delta - \mathrm{B}\Gamma = 1$, every fraction between $\frac{\mathrm{A}}{\Gamma}$ and $\frac{\mathrm{B}}{\Delta}$ has its denominator strictly larger than both Γ and Δ.

Suppose that $\frac{\alpha}{\gamma}$ and $\frac{\beta}{\delta}$ (which have the same sign) have the same sign as a. Then, $L := \frac{-b+\sqrt{D}}{2a}$ lies between them.

Suppose (furthermore) that $\frac{\alpha}{\gamma}$ lies strictly between $\frac{\alpha_i}{\gamma_i}$ and $\frac{\alpha_{i+2}}{\gamma_{i+2}}$. Then $\frac{\beta}{\delta}$ lies on the other side of L as both $\frac{\alpha_i}{\gamma_i}$ and $\frac{\alpha_{i+2}}{\gamma_{i+2}}$, so it must lie either between L and $\frac{\alpha_{i+1}}{\gamma_{i+1}}$, or outside of this

interval, closer to $\frac{\alpha_{i+1}}{\gamma_{i+1}}$ than to L, unless it is exactly equal to $\frac{\alpha_{i+1}}{\gamma_{i+1}}$. If it is the first case, in particular $\frac{\beta}{\delta}$ lies between $\frac{\alpha_{i+2}}{\gamma_{i+2}}$ and $\frac{\alpha_{i+1}}{\gamma_{i+1}}$, with $|\alpha_{i+2}\gamma_{i+1} - \gamma_{i+2}\alpha_{i+1}| = |\beta_{i+1}\gamma_{i+1} - \delta_{i+1}\alpha_{i+1}| = 1$, so the denominator $|\delta|$ is greater than $|\gamma_{i+2}|$. The same argument, applied to the interval with endpoints $\frac{\alpha}{\gamma}$ and $\frac{\beta}{\delta}$, which contains $\frac{\alpha_{i+2}}{\gamma_{i+2}}$, shows that $|\gamma_{i+2}|$ is greater than the $|\delta|$, which is a contradiction. A similar contradiction is met if $\frac{\beta}{\delta}$ lies outside of the interval between L and $\frac{\alpha_{i+1}}{\gamma_{i+1}}$ or if $\frac{\alpha}{\gamma}$ is greater or smaller than every term of $\left(\frac{\alpha_i}{\gamma_i}\right)_{i\geq 1}$. Therefore, we conclude that either $\frac{\alpha}{\gamma}$ or $\frac{\beta}{\delta}$ is equal to $\frac{\alpha_i}{\gamma_i}$ for some i.

Suppose, $\frac{\alpha}{\gamma} = \frac{\alpha_i}{\gamma_i}$. Then, we will start with:

$$\begin{cases} a_i &= a\alpha_i^2 + b\alpha_i\gamma_i + c\gamma_i^2 \\ a' &= a\alpha^2 + b\alpha\gamma + c\gamma^2 \\ \\ b_i &= 2a\alpha_i\beta_i + b(\alpha_i\delta_i + \beta_i\gamma_i) + 2c\gamma_i\delta_i \\ b' &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta \\ \\ c_i &= a\gamma_i^2 + b\gamma_i\delta_i + c\delta_i^2 \\ c' &= a\gamma^2 + b\gamma\delta + c\delta^2 \end{cases}$$

Since $\alpha\delta - \beta\gamma = 1 = \alpha_i\delta_i - \beta_i\gamma_i$, both $\frac{\alpha}{\gamma}$ and $\frac{\alpha_i}{\gamma_i}$ are irreducible, which means $\alpha = \pm\alpha_i$, $\gamma = \pm\gamma_i$. Either way implies that $a' = a_i$ using the first two equations. By multiplying the third equation by $1 = \alpha\delta - \beta\gamma$ and the fourth one by $1 = \alpha_i\delta_i - \beta_i\gamma_i$ and subtracting, then adding $0 = 2b\alpha\beta\gamma_i\delta_i - 2b\alpha\beta\gamma_i\delta_i$ to the right hand side, the expression factors into:

$$\begin{aligned} b_i - b' &= (\beta\delta_i - \beta_i\delta)(2a\alpha\alpha_i + 2b\alpha\gamma_i + 2c\gamma\gamma_i) - (\alpha\gamma_i - \alpha_i\gamma)(2a\beta\beta_i + 2b\beta\delta_i + 2c\delta\delta_i) = \\ &= (\beta\delta_i - \beta_i\delta)(2a\alpha\alpha_i + 2b\alpha\gamma_i + 2c\gamma\gamma_i) = \\ &= \pm(\beta\delta_i - \beta_i\delta)(2a\alpha^2 + 2b\alpha\gamma + 2c\gamma^2) = \\ &= \pm 2a_i(\beta\delta_i - \beta_i\delta) \end{aligned}$$

So we have $b_i \equiv b' \mod 2a_i$. This implies that both f_i and g are half-reduced successors of f_{i-1} and, therefore, they are equal.

If it is $\frac{\beta}{\delta} = \frac{\alpha_i}{\gamma_i} = \frac{\beta_{i-1}}{\delta_{i-1}}$ instead, mutatis mutandis one can conclude that both f_{i-1} and g are half-reduced predecessors of f_i and, therefore, they are equal.

It remains to study the case where $\frac{\alpha}{\gamma}$ and $\frac{\beta}{\delta}$ have an opposite sign to that of a. The procedure is similar to the one presented (and so, it will not be presented here) and the conclusion is that $g = f_{-i}$ for some $i \geq 1$. □

This theorem allows us to compute whether two indefinite forms are equivalent/properly equivalent:

Corollary 3.21. *Let $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = a'x^2 + b'xy + c'y^2$ be two indefinite forms of discriminant D. Let $Ax^2 + By^2 + Cy^2$ be a half-reduced form equivalent to*

f. Then, f and g are properly equivalent if, applying the algorithm to g we eventually get to $Ax^2 + By^2 + Cy^2$. f and g are equivalent if, applying the algorithm to g we eventually get to $Ax^2 + By^2 + Cy^2$ or to $Cx^2 + By^2 + Ay^2$.

Proof. There is nothing more to say about the first statement. The second one is true since $Ax^2 + By^2 + Cy^2$ and $Cx^2 + By^2 + Ay^2$ are (improperly) equivalent via the substitution $(x, y) \mapsto (y, x)$, and both are half-reduced. □

Example 3.22. As we did with definite forms, we can find $h(D)$, the number of proper equivalence classes of primitive forms of discriminant D, for any given D:

D	$h(D)$	D	$h(D)$
5	1	40	2
8	1	41	1
12	1	44	1
13	1	45	2
17	1	48	2
20	1	52	1
21	1	53	1
24	1	56	1
28	1	57	1
29	1	60	2
32	2	61	1
33	1	65	2
37	1	68	1

Before going into degenerate forms, let's see a strong connection between the continued fraction and indefinite forms. But first, we need to know the following results about continued fractions:

Proposition 3.23. *Let $(x_i)_{i\geq 1}$ be a sequence of positive numbers and define the sequences $(p_i)_{i\geq -1}$ and $(q_i)_{i\geq -1}$ as follows:*

$$\begin{aligned} p_{-1} &= 1 \\ p_0 &= 0 \\ p_{i+2} &= x_{i+2}p_{i+1} + p_i \qquad \forall i \geq -1 \end{aligned}$$

$$\begin{aligned} q_{-1} &= 0 \\ q_0 &= 1 \\ q_{i+2} &= x_{i+2}q_{i+1} + q_i \qquad \forall i \geq -1 \end{aligned}$$

Then we have that

$$\frac{p_i}{q_i} = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots + \cfrac{1}{x_i}}}} \qquad \forall i \geq 1$$

Proof. Let's use induction. For $i = 1$ it holds. Suppose it is true for some i and every choice of x_j's. Then,

$$\cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots + \cfrac{1}{\left(x_i + \cfrac{1}{x_{i+1}}\right)}}}} = \frac{\left(x_i + \cfrac{1}{x_{i+1}}\right) p_{i-1} + p_{i-2}}{\left(x_i + \cfrac{1}{x_{i+1}}\right) q_{i-1} + q_{i-2}} =$$

$$= \frac{p_i + \cfrac{1}{x_{i+1}} p_{i-1}}{q_i + \cfrac{1}{x_{i+1}} q_{i-1}} = \frac{x_{i+1} p_i + p_{i-1}}{x_{i+1} q_i + q_{i-1}} = \frac{p_{i+1}}{q_{i+1}}$$

So it is also true for $i + 1$. □

Lemma 3.24. *The continued fractions of a certain number L and $-L$ have the same tail.*

Proof. If L is rational, then both continued fractions are finite, so there is no tail for either of them. Suppose then that L is irrational.

Let the continued fraction of L be

$$x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}}}$$

where x_i is positive for $i \geq 1$, but x_0 not necessarily, then the continued fraction of $-L$ is, as one can check,

$$(-x_0 - 1) + \cfrac{1}{1 + \cfrac{1}{(x_1 - 1) + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}}}},$$

where it is understood that the fraction collapses into

$$(-x_0-1)+\cfrac{1}{(1+x_2)+\cfrac{1}{x_3+\cfrac{1}{x_4+\cfrac{1}{\ddots}}}}$$

when $x_1=1$. The tails of the continued fractions of L and $-L$ coincide. □

Proposition 3.25. *Using the notation formerly presented, the continued fraction of* $\frac{-b_0+\sqrt{D}}{2|a_0|}$ *is*

$$\cfrac{1}{|l_1|+\cfrac{1}{|l_2|+\cfrac{1}{|l_3|+\cfrac{1}{|l_4|+\cfrac{1}{\ddots}}}}}$$

Proof. Let $x_i=|l_i|>0$ for $i\geq 1$. Let $p_i=|\beta_i|$ and $q_i=|\delta_i|$. Then, the conditions of the proposition hold because we already know that $|\beta_{-1}|=|\delta_0|=1$, that $|\beta_0|=|\delta_{-1}|=0$, that $\beta_{i+2}=\beta_{i+1}l_{i+2}-\beta_i$, that $\delta_{i+2}=\delta_{i+1}l_{i+2}-\delta_i$ and that the absolute values of those differences turn out to be the sum of the magnitudes of both terms (thanks to our study of signs in Lemma 3.17).

Therefore, the whole continued fraction is the limit of the truncated continued fractions and so, is the limit of the sequence $\left(\left|\frac{\beta_i}{\delta_i}\right|\right)_{i\geq 1}$, which is $\left|\frac{-b_0+\sqrt{D}}{2a_0}\right|=\frac{-b_0+\sqrt{D}}{2|a_0|}$ □

This proof seems to be pretty direct because we had studied thoroughly the transformations between a half-reduced form and its successor, previously. We will present now a more natural proof of the same result, but unlike the first one, the second will not give an insight to the partial fractions obtained by truncating the continued fractions.

Proof. Let

$$0<\frac{-b_0+\sqrt{D}}{2|a_0|}=\cfrac{1}{x_1+\cfrac{1}{x_2+\cfrac{1}{x_3+\cfrac{1}{x_4+\cfrac{1}{\ddots}}}}}<1$$

be the continued fraction. We need to show that $x_i=|l_i|$ for all i. Then

$$\frac{-b_1+\sqrt{D}}{2|a_1|}=\frac{b_0-2l_1c_0+\sqrt{D}}{2|c_0|}=\frac{b_0-2|l_1||c_0|+\sqrt{D}}{2|c_0|}=\frac{b_0+\sqrt{D}}{2|c_0|}-|l_1|=-|l_1|+\frac{2|a_0|}{-b_0+\sqrt{D}}=$$

$$= -|l_1| + x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}},$$

where we used that l_1 and c_0 have the same sign, as we discovered in Lemma 3.17.

The fact that f_1 is half-reduced implies that $0 < \frac{-b_1+\sqrt{D}}{2|a_1|} < 1$, so we have $x_1 = |l_1|$. Then,

$$\frac{-b_1 + \sqrt{D}}{2|a_1|} = \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{x_5 + \cfrac{1}{\ddots}}}}}$$

By repeating the argument, we prove $x_2 = |l_2|$, $x_3 = |l_3|$, and so on. □

Theorem 3.26. *For each $g(x, y) = a'x^2 + b'xy + c'y^2$ properly equivalent to a given indefinite form $f(x, y) = ax^2 + bxy + cy^2$ of discriminant D, the continued fraction of $\frac{-b'+\sqrt{D}}{2a'}$ is, from a point onwards, the same as the continued fraction of $\frac{-b+\sqrt{D}}{2a}$.*

Proof. Consider a sequence of forms, starting with f and ending with a half-reduced form h, such that each form is the successor of the previous one. This can be done, for example using the algorithm. Similarly, a similar sequence exists, but starting from g instead of f. If we could prove that the tail of the continued fractions is the same between neighbours, we will have that f and h, g and h and therefore f and g will also have the same tails, thus proving the theorem. So, we may suppose that f and g are neighbours: $g(x, y) = a'x^2 + b'xy + c'y^2 = cx^2 + (-b + 2lc)xy + (l^2c - lb + a)y^2$

Similar to what we did before,

$$\frac{-b' + \sqrt{D}}{2a'} = \frac{b - 2lc + \sqrt{D}}{2c} = \frac{b + \sqrt{D}}{2c} - l = -l - \frac{2a}{-b + \sqrt{D}}$$

We have that if

$$\frac{-b' + \sqrt{D}}{2a'} = x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}}},$$

then

$$\frac{-b'+\sqrt{D}}{2a'} = -l - \frac{2a}{-b+\sqrt{D}} = -l - \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}}}}.$$

If x_0 is non-negative, then the negative sign in front of the fraction and the addition of an integer $(-l)$ do not affect the tail of the continued fraction and we are done.

If $x_0 < 0$, then the previous quantity equals

$$-l + \cfrac{1}{(-x_0 - 1) + \cfrac{1}{1 + \cfrac{1}{(x_1 - 1) + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cfrac{1}{\ddots}}}}}}},$$

which is a continued fraction and it has the same tail. □

Corollary 3.27. *Given an indefinite form $f(x,y) = ax^2 + bxy + cy^2$ of discriminant D, the continued fraction of $\frac{-b+\sqrt{D}}{2a}$ is k-periodic from one point onwards, being k the number of half-reduced forms properly equivalent to f.*

Proof. It follows directly from the last two results and the fact that the sequence $(l_i)_{i\in\mathbb{Z}}$ is k-periodic. □

4 Degenerate forms

As we did with definite forms, we will present some set of 'reduced' degenerate forms and we will prove that there are not proper equivalences among them and that every degenerate form is properly equivalent to one of them:

Definition 4.1. A degenerate form $f(x, y) = ax^2 + bxy + cy^2$ of discriminant D is called *(left)-diminished* if $a = 0$, $b \geq 0$ and $0 \leq c < b$ whenever $b \neq 0$.

Again, it is common to call diminished forms 'reduced' due to its similarities with reduced definite forms. Instead, I chose to use the non-standard term 'diminished' to design the reduced notion for degenerate forms. The prefix 'left' comes from the fact that $a = 0$. We could have defined instead the notion of 'right-diminished' by changing the roles of a and c, and all the following would be analogous. If the term 'diminished' is used without specifying left or right, one should understand it is left-diminished.

Theorem 4.2. *Each degenerate form is properly equivalent to a unique diminished form.*

Proof. This proof follows Gauß's in [1, §§206-207].

Given a degenerate form f, we need to find a diminished form properly equivalent to it. Provided that f is degenerate, we may write $f(x, y) = (Ax + By)(\Gamma x + \Delta y)$. Then its discriminant is $D = (A\Delta - B\Gamma)^2$. Let's denote $d = A\Delta - B\Gamma$ which, without loss of generality, we may suppose to be non-negative. Otherwise, just change the order of the factors.

Now, move the content to the first factor, so that $\gcd(\Gamma, \Delta) = 1$. Let $\alpha = -\Delta$ and $\gamma = \Gamma$. Since α and γ are relatively prime, we can find β and δ such that $\alpha\delta - \beta\gamma = -\Delta\delta - \beta\Gamma = 1$. Let $g(x, y) = f(\alpha x + \beta y, \gamma x + \delta y) = a'x^2 + b'xy + c'y^2$.

We then have

$$a' = f(\alpha, \gamma) = (A(-\Delta) + B\Gamma)(\Gamma(-\Delta) + \Delta\Gamma) = 0$$

$$\begin{aligned} b' &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta = \\ &= 2A\Gamma(-\Delta)\beta + (B\Gamma + A\Delta)(-\Delta\delta + \beta\Gamma) + 2B\Delta\Gamma\delta = \\ &= A\Delta(-\Delta\delta - \beta\Gamma) - B\Gamma(-\Delta\delta - \beta\Gamma) = \\ &= d \geq 0 \end{aligned}$$

If $d = 0$ or $0 \leq c' < d$, we are done. Otherwise, let l be the integer such that $0 \leq c' + ld < d$ and make the substitution $(x, y) \mapsto (x + ly, y)$. The form $g(x, y) = dxy + c'xy$ is transformed into $h(x, y) = dxy + (c' + ld)y^2$. The form h is the diminished form we were looking for. Since all transformations used to arrive here were proper, h is properly equivalent to f.

What remains now is to prove that two different diminished forms are not properly equivalent.

Suppose $f(x,y) = bxy + cy^2$ and $g(x,y) = b'xy + c'y^2$ are diminished and properly equivalent. Then, they have the same discriminant $b'^2 = b'^2 - 4a'c' = D = b^2 - 4ac = b^2$ and so, since $b, b' \geq 0$, we have $b = b' =: d$.

Let $g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$, then

$$\begin{cases} c' &= f(\beta,\delta) = d\beta\delta + c\delta^2 \\ d &= d(\alpha\delta + \beta\gamma) + 2c\gamma\delta = d + 2d\beta\gamma + 2c\gamma\delta = d + 2\gamma(d\beta + c\delta) \\ 0 &= f(\alpha,\gamma) = d\alpha\gamma + c\gamma^2 = \gamma(d\alpha + c\gamma) \end{cases}$$

From the second equation, we get that either $\gamma = 0$ or $d\beta + c\delta = 0$. From the third equation, $\gamma = 0$ or $d\alpha + c\gamma = 0$. If $\gamma \neq 0$, then $0 = \alpha(d\beta + c\delta) - \beta(d\alpha + c\gamma) = c$ and $0 = \gamma(d\beta + c\delta) - \delta(d\alpha + c\gamma) = -d$, so f is the zero form and $g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$ as well. So, we may suppose $\gamma = 0$ so $\alpha = \delta = \pm 1$ and, looking at the first equation, we get $c' = c \pm d\beta$, which means that c and c' are congruent modulo d. If $d = 0$, $c = c'$ so $f = g$ and we are done. Otherwise, in order to be diminished, both c and c' must lie between 0 and $d-1$ inclusive, which forces them to be equal. Therefore, we also have $f = g$.

We have proved existence and uniqueness, so we are done. □

Corollary 4.3. *If $d > 0$, there are d proper equivalence classes of a given discriminant $D = d^2$, $\varphi(d)$ of which are primitive (φ is Euler's totient function).*

There is a proper equivalence class of forms of discriminant 0 for each integer M, but only two of them are primitive.

Proof. It is sufficient to count how many diminished forms satisfy the conditions. For the first part, we see that there are exactly d of such forms, namely $dxy + cy^2$ for $c = 0, 1, 2, \ldots, d-1$, and $\mathrm{Cont}(dxy + cy^2) = \gcd(d, c)$, which equals 1 in exactly $\varphi(d)$ of the cases, by definition of φ.

The diminished forms of discriminant 0 are those of the form My^2, being M an integer. Only if $M = \pm 1$, the form is primitive. □

The set of diminished forms of a given discriminant $D = d^2$ is therefore a set of representatives of the proper equivalence classes. Note that the set of right-diminished forms (i.e. those of the form $ax^2 + dxy$, assuming $d \leq 0$, satisfying $0 \leq a < d$ whenever $d > 0$) is analogously also a set of representatives of the classes.

Corollary 4.4. *If $f(x,y) = ax^2 + bxy + cy^2$ has discriminant 0, then it only represents non-negative or non-positive numbers.*

Proof. Let $h(x,y)$ be a diminished form properly equivalent to f. Then, $h(x,y) = My^2$ for some integer M. The sign of M determines the sign of $h(x_0, y_0)$ for every choice of x_0

and y_0, therefore h always represents non-negative or non-positive numbers. Since f and h are equivalent, also does f. $\square$

We see a resemblance between forms of discriminant 0 and definite forms.

Definition 4.5. We say that a form f of discriminant 0 is positive semi-definite if it represents only non-negative numbers.
We say it is negative semi-definite if it represents only non-positive numbers.

Theorem 4.6. *Suppose both $f(x,y) = bxy + cy^2$ and $g(x,y) = b'xy + c'y^2$ are diminished and have the same discriminant d^2 and content n. Then they are improperly equivalent if and only if $\frac{c}{n} \cdot \frac{c'}{n} \equiv 1 \mod \frac{d}{n}$.*

Proof. First of all, we have that $b'^2 = b'^2 - 4a'c' = d^2 = b^2 - 4ac = b^2$ and so, since $b, b' \geq 0$, we have $b = b' = d$ (we just chose the sign of d to be non-negative), and so n divides d and the congruence makes sense.

The last condition is equivalent to $cc' \equiv n^2 \mod nd$, so we will work using this one instead. Let $g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$, then

$$\begin{cases} c' &= f(\beta,\delta) = d\beta\delta + c\delta^2 \\ d &= d(\alpha\delta + \beta\gamma) + 2c\gamma\delta = d + 2d\alpha\delta + 2c\gamma\delta = d + 2\delta(d\alpha + c\gamma) \\ 0 &= f(\alpha,\gamma) = d\alpha\gamma + c\gamma^2 = \gamma(d\alpha + c\gamma) \end{cases}$$

Since it cannot be that both γ and δ are zero, we have that $d\alpha + c\gamma = 0$. Therefore, the ratio $\alpha : \gamma$ equals the ratio $-c : d$. Since $\alpha\delta - \beta\gamma = -1$, α and γ are relatively prime and we can decide the sign of n such that $-c = n\alpha, d = n\gamma$. Then, the first equation reads $c' = d\beta\delta + c\delta^2 = -\delta n(\alpha\delta - \beta\gamma) = \delta n$, so $cc' = \delta nc = -n^2\alpha\delta = n^2 - n^2\beta\gamma = n^2 - \beta nd$ and the congruence is true.

Conversely, Let $\alpha = -\frac{c}{n}$, $\beta = \frac{n^2 - cc'}{nd}$, $\gamma = \frac{d}{n}$ and $\delta = \frac{c'}{n}$, which are integers that satisfy $\alpha\delta - \beta\gamma = -\frac{cc'}{n^2} - \frac{n^2 - cc'}{n^2} = -1$. One can check that

$$\begin{aligned} c' &= \frac{(n^2 - cc')c'}{n^2} + c\frac{(c')^2}{n^2} = d\beta\delta + c\delta^2 \\ d &= d + 2\frac{c'}{n}\left(-d\frac{c}{n} + c\frac{d}{n}\right) = d + 2\delta(d\alpha + c\gamma) \\ 0 &= 2\frac{d}{n}\left(-d\frac{c}{n} + c\frac{d}{n}\right) = 2\gamma(d\alpha + c\gamma) \end{aligned}$$

So $g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$ and $f(x,y)$ are improperly equivalent. $\square$

5 Automorphisms of forms

Definition 5.1. Both a transformation $(x, y) \mapsto (\alpha x + \beta y, \gamma x + \delta y)$ and the associated matrix $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$ are called an *automorphism of* f if f is invariant under it:

$$f(x, y) = f(\alpha x + \beta y, \gamma x + \delta y)$$

The automorphism U is called *proper/improper* if the transformation was so, in other words, if $\det U$ equals 1 or -1, respectively. We say that the transformations $(x, y) \mapsto (x, y)$ and $(x, y) \mapsto (-x, -y)$ are *trivial automorphisms*.

The name 'trivial' in the former definition comes from the fact that they are automorphisms of all forms. One can see that the set of automorphisms and the set of proper automorphisms of a given form has a subgroup structure of $\mathrm{GL}_2(\mathbb{Z})$ and $\mathrm{SL}_2(\mathbb{Z})$, respectively. In terms of the action of these groups over the set of forms, the group of automorphisms of f is the stabilizer of f.

Definition 5.2. Forms for which improper automorphisms exist are called *ambiguous*.

Note that ambiguous forms are precisely those who are improperly equivalent to themselves, and those whose equivalence and proper equivalence class coincide.

If $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = a'x^2 + b'xy + c'y^2$ are equivalent via the matrix A, which means

$$A^\top \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} A = \begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix}$$

and U is an automorphism of g, then one can check that $A^{-1}UA$ is an automorphism of f. Additionally, the automorphisms of $-f$ and f coincide, so we may restrict ourselves to study the automorphisms of reduced, half-reduced and diminished forms.

Theorem 5.3. *Let $f(x, y) = ax^2 + bxy + cy^2$ be a reduced form. Then if $f(x, y)$ does not equal $g(x, y) = ax^2 + axy + ay^2$ nor $h(x, y) = ax^2 + ay^2$, then f only has trivial proper automorphisms. g has additional proper automorphisms*

$$\begin{pmatrix} 0 & \pm 1 \\ \mp 1 & \mp 1 \end{pmatrix}, \begin{pmatrix} \pm 1 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}$$

and h has additional proper automorphisms

$$\begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}$$

Proof. Let $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ be a proper automorphism of f. Then we have

$$\begin{cases} a &= a\alpha^2 + b\alpha\gamma + c\gamma^2 = f(\alpha, \gamma) \\ b &= 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta \\ c &= a\beta^2 + b\beta\delta + c\delta^2 = f(\beta, \delta) \end{cases}$$

Since f is reduced, we have $a = f(\alpha, \gamma) \geq (a - |b| + c)\min\{\alpha^2, \gamma^2\} \geq a\min\{\alpha^2, \gamma^2\}$. Similarly, $c \geq c\min\{\beta^2, \delta^2\}$. Therefore, α, β, γ and δ can only take the values $-1, 0, 1$. Since $\alpha\delta - \beta\gamma = \pm 1$, one of them must be equal to 0.

If $\alpha = 0$, then $\beta\gamma = -(\alpha\delta - \beta\gamma) = -1$, so $\beta = -\gamma = \pm 1$ and $a = f(0, \gamma) = c\gamma^2 = c$. This implies that $b \geq 0$. $b = 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta = -b + 2c\gamma\delta$, so $2c\gamma\delta$ is also non-negative. This means that either $\delta = 0, b = 0$ and

$$f(x, y) = ax^2 + ay^2, \quad U = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}$$

or $\delta = \gamma, b = c$ and

$$f(x, y) = ax^2 + axy + ay^2, \quad U = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & \mp 1 \end{pmatrix}$$

If $\beta = 0$, then $\alpha\delta = 1$ and $b = 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta = b + 2c\gamma\delta$, so $\gamma = 0$ and

$$f(x, y) = ax^2 + bxy + cy^2, \quad U = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

Similarly, if $\gamma = 0$ we have

$$f(x, y) = ax^2 + bxy + cy^2, \quad U = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

and if $\delta = 0$, we have either

$$f(x, y) = ax^2 + ay^2, \quad U = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}$$

or

$$f(x, y) = ax^2 + axy + ay^2, \quad U = \begin{pmatrix} \pm 1 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}$$

□

Corollary 5.4. *If f is definite, the group of proper automorphisms of f is isomorphic to $\mathbb{Z}/6\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z}$, depending on whether its reduced properly equivalent form equals $ax^2+axy+ay^2$, $ax^2 + ay^2$ or neither, being $a = \pm Cont(f)$.*

Proof. By conjugation, the proper automorphism group of f is isomorphic to the proper automorphism group of a reduced form g in its proper equivalence class. The result follows from studying the powers of $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. □

Theorem 5.5. *Let $f(x, y) = ax^2 + bxy + cy^2$ be a half-reduced form. There exists a matrix A such that all proper automorphisms of f are of the form $\pm A^n$ for some $n \in \mathbb{Z}$.*

Proof. Let $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ be a proper automorphism of f. Then, f is transformed via U into another half-reduced form g (namely, f itself). Looking at the proof of Theorem 3.13, we can conclude that there exists an integer i such that $g = f_i$ and $\pm U = \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix}$. Since $f = g = f_i$, we know that i is an integer multiple of the length k of the cycle. Let's define

$$A := \begin{pmatrix} 0 & -1 \\ 1 & l_1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & -1 \\ 1 & l_k \end{pmatrix}$$

Since $(l_i)_{i\in\mathbb{Z}}$ is k-periodic, we have that either

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha_0 & \beta_0 \\ \gamma_0 & \delta_0 \end{pmatrix} = \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_{i+1} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_{i+2} \end{pmatrix} \cdots \begin{pmatrix} 0 & -1 \\ 1 & l_0 \end{pmatrix} = \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} A^{\frac{i}{k}}$$

or

$$\begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} = \begin{pmatrix} \alpha_0 & \beta_0 \\ \gamma_0 & \delta_0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & l_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & -1 \\ 1 & l_i \end{pmatrix} = \begin{pmatrix} \alpha_0 & \beta_0 \\ \gamma_0 & \delta_0 \end{pmatrix} A^{\frac{i}{k}} = A^{\frac{i}{k}},$$

depending on the sign of i. In either case, we obtain the desired result. □

Corollary 5.6. *If f is indefinite, the group of proper automorphisms of f is isomorphic to $\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$.*

Proof. By conjugation, the automorphism group of f is isomorphic to the automorphism group of a half-reduced form g in its proper equivalence class. The result follows from the fact that A^n has greater and greater coefficients (in absolute value) as $|n|$ grows, due to Lemma 3.17, and so A does not have a finite order. □

Theorem 5.7. *Let $f(x, y) = dxy + cy^2$ be a half-reduced form. If $d \neq 0$, then f only has trivial proper automorphisms. If $d = 0 \neq c$, the proper automorphisms of f are precisely*

$$\left\{ \begin{pmatrix} \pm 1 & \beta \\ 0 & \pm 1 \end{pmatrix} \right\}_{\beta\in\mathbb{Z}}.$$

If $d = 0 = c$, then any matrix in $\mathrm{SL}_2(\mathbb{Z})$ is a proper automorphism.

Proof. We will proceed very similarly to the uniqueness part of the proof of Theorem 4.2.

Let $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ be a proper automorphism of f. Then we have

$$\begin{cases} c &= f(\beta, \delta) = d\beta\delta + c\delta^2 \\ d &= d(\alpha\delta + \beta\gamma) + 2c\gamma\delta = d + 2d\beta\gamma + 2c\gamma\delta = d + 2\gamma(d\beta + c\delta) \\ 0 &= f(\alpha, \gamma) = d\alpha\gamma + c\gamma^2 = \gamma(d\alpha + c\gamma) \end{cases}$$

From the second equation, we get that either $\gamma = 0$ or $d\beta + c\delta = 0$. From the third equation, $\gamma = 0$ or $d\alpha + c\gamma = 0$. If $\gamma \neq 0$, then $0 = \alpha(d\beta + c\delta) - \beta(d\alpha + c\gamma) = c$ and

$0 = \gamma(d\beta + c\delta) - \delta(d\alpha + c\gamma) = -d$, so f is the zero form, whose group of automorphisms is the whole $\mathrm{SL}_2(\mathbb{Z})$.

Otherwise, we have $\gamma = 0$ so $\alpha = \delta = \pm 1$ and, looking at the first equation, we get $c = c \pm d\beta$, which means that either $\beta = 0$ and

$$f(x, y) = dxy + cy^2, \quad U = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

or $d = 0$ and

$$f(x, y) = cy^2, \quad U = \begin{pmatrix} \pm 1 & \beta \\ 0 & \pm 1 \end{pmatrix}$$

□

Corollary 5.8. *If f is degenerate and its discriminant is non-zero, then f only has trivial automorphisms. If its discriminant is zero, then its group of automorphisms is isomorphic to $\mathrm{SL}_2(\mathbb{Z})$ or $\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$, depending on whether f is the zero form or not.*

Proof. By conjugation, the automorphism group of f is isomorphic to the automorphism group of a diminished form $g(x, y) = dxy + cy^2$ in its proper equivalence class. The condition about the discriminant being zero or not is equivalent to whether $d = 0$ or not. The result follows since

$$\begin{pmatrix} \pm 1 & \beta \\ 0 & \pm 1 \end{pmatrix} = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\pm\beta}$$

and this matrix does not have a finite order. □

One corollary of the study of automorphisms is the following theorem:

Theorem 5.9. $\mathrm{SL}_2(\mathbb{Z})$ *is generated by the matrices* $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Proof. Consider the form $g(x, y) = 2x^2 + 2xy + 3y^2$, which is positive definite and reduced. We know that its group of automorphisms is $\{\mathrm{Id}, -\mathrm{Id}\}$. Given any matrix A in $\mathrm{SL}_2(\mathbb{Z})$, consider the result of acting A on g, which is a form f properly equivalent to g. Apply the algorithm to f until we reach g. Since the algorithm only applies transformations S and T^k, we see that there is a transformation $B \in \langle S, T \rangle$ such that acting B on f returns g. Therefore, AB is an automorphism of g and therefore $AB = \pm \mathrm{Id}$. So, $A = \pm \mathrm{Id} \cdot B^{-1} \in \langle S, T \rangle$ because $-\mathrm{Id} = S^2$, which means $\langle S, T \rangle = \mathrm{SL}_2(\mathbb{Z})$.

We used the form $2x^2 + 2xy + 3y^2$, but there was nothing special with that particular form. In fact, all definite or indefinite forms would have sufficed to prove the result since one can check that their automorphisms lie in $\langle S, T \rangle$ and we know that the algorithms only make use of S, T^k and $\begin{pmatrix} 0 & -1 \\ 1 & l \end{pmatrix} = S \cdot T^l$. □

However, the theory of binary quadratic forms is not required to prove this theorem. By left-multiplying by T^k and S a given matrix $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$, and observing what happens to the first and third coefficient in each case, one can derive a much simpler proof using an argument similar to Euclid's algorithm's.

The rest of this section takes ideas from [5, pp. 27-29]. Automorphisms have a close connection to the solutions of a particular Pell equation, but to see the connection, we need to prove the following lemma:

Lemma 5.10. *Let $f(x,y) = ax^2+bxy+cy^2$ and $g(x,y) = a'x^2+b'xy+c'y^2 = f(\alpha x+\beta y, \gamma x+\delta y)$ be two equivalent forms of discriminant D. Let $\omega = \frac{-b+\sqrt{D}}{2a}, \omega' = \frac{-b'+\sqrt{D}}{2a'} \in \mathbb{C}$, taking the square root to be either positive or positive imaginary. Then, it happens that $\omega = \frac{\alpha\omega'+\beta}{\gamma\omega'+\delta}$*

Proof. A simple computation shows that the result is true. □

Theorem 5.11. *Let $f(x,y) = ax^2+bxy+cy^2$ be a primitive form of discriminant D. There is a one-to-one correspondence between the proper automorphisms of f and the solutions to the equation $x^2-Dy^2 = 4$, and, if D is not a square, there is a one-to-one correspondence between the improper automorphisms of f and the solutions to the equation $x^2-Dy^2 = -4$*

Proof. We will prove both results all together. Each time $\pm$ is used, the $+$ sign is for the first result and the $-$ sign for the second one.

First, suppose f is non-degenerate. We will prove that the following mapping is one-to-one: given a solution (X,Y), we construct the automorphism

$$U(X,Y) = \begin{pmatrix}\frac{X-Yb}{2} & -cY\\ aY & \frac{X+Yb}{2}\end{pmatrix}$$

First, we note that the determinant of $U(X,Y)$ is $\frac{X-Yb}{2}\cdot\frac{X+Yb}{2} + acY^2 = \frac{X^2-DY^2}{4} = \pm 1$. The coefficients of $U(X,Y)$ have to be integers since $X-Yb$ and $X+Yb$ have the same parity and the determinant is an integer. Therefore, we have that $U(X,Y) \in \mathrm{GL}_2(\mathbb{Z})$.

Let $g(x,y) = a'x^2+b'xy+c'y^2$ the result of applying $U(X,Y)$ to f. By applying the formulas in (1.3), which relate the coefficients of f and g, we obtain that $a' = a$, $b' = b$ and $c' = c$ and therefore $U(X,Y)$ is an automorphism.

The injectivity of the mapping comes from the fact that X is completely determined since it is the trace of $U(X,Y)$; and since $\gcd(a,b,c) = 1$, we have that at least one of the coefficients is non-zero and therefore Y is determined by either the second coefficient, the third one or the difference between the first and the fourth.

To prove that the mapping is surjective, consider $\omega = \frac{-b+\sqrt{D}}{2a} \in \mathbb{C}$, where the square root is chosen either positive or positive imaginary. Let $U = \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ be an automorphism of f.

Then, we have that $\omega = \frac{\alpha\omega+\beta}{\gamma\omega+\delta}$, which implies $\gamma\omega^2 + (\delta-\alpha)\omega - \beta = 0$. Since f is not degenerate, there is only one quadratic equation that ω can satisfy (except for scalar multiplication), and ω already satisfies $a\omega^2 + b\omega + c = 0$. So, because f is primitive, there exists an integer Y such that $\gamma = aY$, $-\beta = cY$ and $\delta - \alpha = bY$. Then, calling $X = \alpha + \delta \in \mathbb{Z}$, we can write $U = U(X,Y) = \begin{pmatrix} \frac{X-Yb}{2} & -cY \\ aY & \frac{X+Yb}{2} \end{pmatrix}$. The solution to the equation we were looking for is (X,Y), which is indeed a solution because $\pm 1 = \alpha\delta - \beta\gamma = \frac{X-Yb}{2} \cdot \frac{X+Yb}{2} + acY^2 = \frac{X^2-DY^2}{4}$.

Therefore, the mapping is one-to-one.

It remains to show the first result for degenerate forms. The fact that $(\gamma : \delta - \alpha : -\beta) = (a : b : c)$ does not need to hold anymore, so we must use another approach.

For the positive equation, if the form has positive discriminant, since the only two square numbers differing by four are 4 and 0, the equation $X^2 - d^2Y^2 = 4$ has $(2,0)$ and $(-2,0)$ as its only solutions, and the only proper automorphisms of such form are the trivial ones. If the form f has discriminant equal to 0, consider the form My^2 properly equivalent to it. f being primitive implies $M = \pm 1$. Then, considering the transformation $(x,y) \mapsto (\alpha x + \beta y, \gamma x + \delta y)$ that turns h into f, we can write $f(x,y) = M(\gamma x + \delta y)^2 = (M\gamma^2)\, x^2 + (2M\gamma\delta)xy + (M\delta^2)\, y^2$. If U is a proper automorphism of g, then

$$U = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \begin{pmatrix} \pm 1 & -MY \\ 0 & \pm 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \pm 1 - M\gamma\delta Y & -M\delta^2 Y \\ M\gamma^2 Y & \pm 1 + M\gamma\delta Y \end{pmatrix} = U(\pm 2, Y)$$

for some integer Y, and the set $\{(\pm 2, Y)\}_{Y\in\mathbb{Z}}$ is precisely the set of solutions of the equation $x^2 - 0y^2 = 4$. □

Although we could have noticed it when we studied proper automorphisms, it is now clearer that the number of automorphisms of a primitive form only depends on its discriminant, rather than their proper class.

Since automorphisms have a group structure, we can translate this structure via the bijection onto the set of solutions to Pell's equation. Given two solutions (X,Y) and (X',Y'), we may construct its composition:

$$\begin{pmatrix} \frac{X-bY}{2} & -cY \\ aY & \frac{X+bY}{2} \end{pmatrix} \begin{pmatrix} \frac{X'-bY'}{2} & -cY' \\ aY' & \frac{X'+bY'}{2} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{XX'-DYY'}{4} - \frac{b}{2}(XY' + X'Y) & -\frac{c}{2}(XY' + X'Y) \\ \frac{a}{2}(XY' + X'Y) & \frac{XX'-DYY'}{4} + \frac{b}{2}(XY' + X'Y) \end{pmatrix}$$

The composition must be $(X,Y) \circ (X',Y') = \left(\frac{XX'-DYY'}{2}, \frac{XX'+YY'}{2}\right)$

We will revisit proper automorphisms when we talk about quadratic fields in the following sections, since they are closely related to the units.

6 Composition

In the sections before, we made a natural definition with respect to the representation of numbers by forms: equivalent forms. However, we also restricted this definition by introducing the notion of proper equivalence. It is true that this concept has been useful to prove the finiteness of equivalence classes and for determining whether two given forms are equivalent or not. Nonetheless, the great success of proper equivalence is the fact that we can give a group structure to the set of properly equivalent primitive classes of a given discriminant.

In fact, the operation of composition was constructed by Legendre before the notion of proper equivalence, but it was a multivalued operation. When Gauß introduced the notion of proper equivalence he stated that «the usefulness of these distinctions will soon be made clear» [1, §158], he was anticipating the great progress it produces concerning composition. Nowadays, we talk about Gaussian composition of forms.

Definition 6.1. Let $f(x,y) = ax^2+bxy+cy^2$, $g(x,y) = a'x^2+b'xy+c'y^2$, and $h(x,y) = a''x^2+b''xy+c''y^2$ be three primitive forms, all of the same discriminant. We say that h is a *composition* of f and g if there exist two bilinear forms $B(x,y;z,w) = pxz+qxw+ryz+syw$ and $B'(x,y;z,w) = p'xz+q'xw+r'yz+s'yw$ having integer coefficients which satisfy that $f(x,y)\cdot\ g(z,w) = h(B(x,y;z,w),B'(x,y;z,w))$, and also $pq'-p'q = a$ and $pr'-p'r = a'$.
In matrix form:

$$\left[\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}a & \frac{b}{2}\\ \frac{b}{2} & c\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\right]\left[\begin{pmatrix}z & w\end{pmatrix}\begin{pmatrix}a' & \frac{b'}{2}\\ \frac{b'}{2} & c'\end{pmatrix}\begin{pmatrix}z\\ w\end{pmatrix}\right] = \begin{pmatrix}B & B'\end{pmatrix}\begin{pmatrix}a'' & \frac{b''}{2}\\ \frac{b''}{2} & c''\end{pmatrix}\begin{pmatrix}B\\ B'\end{pmatrix},$$

$$\text{where } \begin{pmatrix}B\\ B'\end{pmatrix} = \begin{pmatrix}p & q & r & s\\ p' & q' & r' & s'\end{pmatrix}\begin{pmatrix}xz\\ xw\\ yz\\ yw\end{pmatrix} \text{ and } pq'-p'q = a;\ pr'-p'r = a'$$

This definition naturally arises when considering products of representations: if m is represented by f and n is represented by g, then mn will be represented by any composition of f and g.

Gauß called the composition to be direct if the last two equalities held, and indirect otherwise, but we will only consider direct composition and so drop out the term 'direct'. By writing out and expanding $f(x,y)\cdot g(z,w) = h(B(x,y;z,w),B'(x,y;z,w))$ and manipulating the resulting equations, Gauß managed to prove, among other relations, that the following equalities always held. We will omit the procedure (which can be found in [1, §235][1] since it is just uninteresting algebraic manipulation, and we will only present the results:

Define

$$d_{12} = pq'-p'q,\quad d_{13} = pr'-p'r,\quad d_{14} = ps'-p's$$

[1]Gauß defines composition in a much larger sense: he does not restrict the three forms to be primitive nor to have the same discriminant. We restricted the definition in order to make any two appropriate forms composable. The results we present have been adapted to our case.

$$d_{23} = qr' - q'r, \quad d_{24} = qs' - q's, \quad d_{34} = rs' - r's;$$

then,

$$\begin{cases} d_{12}^2 &= a^2 \\ d_{12}d_{34} &= ac \\ d_{12}\left(d_{14} - d_{23}\right) &= ab \\ d_{13}^2 &= a'^2 \\ d_{13}d_{24} &= a'c' \\ d_{13}\left(d_{14} + d_{23}\right) &= a'b' \end{cases} \tag{6.1}$$

The extra two conditions in the definition of composition are therefore only determining the sign of those expressions. An important observation is that, once the sign of d_{12} and d_{13} is determined, the rest of the d_{ij}'s are completely determined by (6.1). In direct composition:

$$\begin{array}{lll} d_{12} = a, & d_{13} = a', & d_{14} = \frac{b+b'}{2} \\ d_{23} = \frac{b'-b}{2}, & d_{24} = c', & d_{34} = c; \end{array} \tag{6.2}$$

The condition of h being primitive has been imposed in the definition, but it can be relaxed, since Gauß's lemma will ensure that.

Note that if a form h is a composition of f and g, then it is also a composition of g and f, as it is sufficient to change all the x's and the y's by z's and w's, and vice versa, which ends up swapping the values of q and r and the values of q' and r', so the two extra conditions are still satisfied.

Let's see what happens if we drop the last two conditions in the definition of composition: let h, h', h'', h''' be defined such that

$$\begin{aligned} f(x,y)\cdot g(z,w) &= h(B(x,y;z,w), B'(x,y;z,w)) \text{ with } d_{12} = a,\ d_{13} = a'; \\ f(x,y)\cdot g(z,w) &= h'(B(x,y;z,w), B'(x,y;z,w)) \text{ with } d_{12} = a,\ d_{13} = -a'; \\ f(x,y)\cdot g(z,w) &= h''(B(x,y;z,w), B'(x,y;z,w)) \text{ with } d_{12} = -a,\ d_{13} = a'; \\ f(x,y)\cdot g(z,w) &= h'''(B(x,y;z,w), B'(x,y;z,w)) \text{ with } d_{12} = -a,\ d_{13} = -a'; \end{aligned}$$

Then we have

$$\begin{aligned} f(-x,y)\cdot g(z,w) &= h'(B(-x,y;z,w), B'(-x,y;z,w)) \text{ with } d_{12} = a,\ d_{13} = a'; \\ f(x,y)\cdot g(-z,w) &= h''(B(x,y;-z,w), B'(x,y;-z,w)) \text{ with } d_{12} = a,\ d_{13} = a'; \\ f(-x,y)\cdot g(-z,w) &= h'''(B(-x,y;-z,w), B'(-x,y;-z,w)) \text{ with } d_{12} = a,\ d_{13} = a'. \end{aligned}$$

In other words, if we change the sign on the extra conditions, we are just making a direct composition of $f(\pm x, y)$ and $g(\pm z, w)$. Since later we will expand the notion of composition to proper classes, and the equivalence class of a form $f(x,y)$ is the union of the proper classes of $f(\pm x, y)$, we see that allowing non-direct composition will end up being an operation on the classes. However, it will not be single-valued: one can see that $h \sim h'''$ and that $h' \sim h''$, but they need not be equivalent altogether, and the equivalences are not necessarily proper.

Direct composition is what will become the group operation at the level of proper classes, so our goal is to prove the following theorem:

Theorem 6.2. *The following statements are true:*

1. *The composition is well defined at the level of classes. This means:*
   - *There always exists a composition of primitive forms f and g of the same discriminant.*
   - *All compositions of f and g belong to the same proper equivalence class.*
   - *The proper class of the composition only depends on the proper classes of f and g, rather than f and g themselves.*
2. *The set of primitive proper equivalence classes having a given discriminant is a finite abelian group with composition as its operation.*

Some examples of composition may be:

Example 6.3. A well-known identity:

$$\left(x^2+y^2\right)\left(z^2+w^2\right)=B^2+B'^2=(xz-yw)^2+(xw+yz)^2,$$

$$\text{with } \begin{pmatrix}B\\B'\end{pmatrix}=\begin{pmatrix}1&0&0&-1\\0&1&1&0\end{pmatrix}\begin{pmatrix}xz\\xw\\yz\\yw\end{pmatrix}$$

Example 6.4. More generally:

$$\left(x^2+ny^2\right)\left(z^2+nw^2\right)=B^2+nB'^2=(xz-nyw)^2+n(xw+yz)^2,$$

$$\text{with } \begin{pmatrix}B\\B'\end{pmatrix}=\begin{pmatrix}1&0&0&-n\\0&1&1&0\end{pmatrix}\begin{pmatrix}xz\\xw\\yz\\yw\end{pmatrix}$$

Example 6.5. A not so well-known identity is the following:

$$\left(x^2+xy+ny^2\right)\left(z^2+zw+nw^2\right)=B^2+BB'+nB'^2=$$

$$=(xz-nyw)^2+(xz-nyw)(xw+yz+yw)+n(xw+yz+yw)^2,$$

$$\text{with } \begin{pmatrix}B\\B'\end{pmatrix}=\begin{pmatrix}1&0&0&-n\\0&1&1&1\end{pmatrix}\begin{pmatrix}xz\\xw\\yz\\yw\end{pmatrix}$$

By letting $n = -\frac{D}{4}$ and $n = \frac{1-D}{4}$, respectively, on Example 6.4 and Example 6.5, we see, if we accept Theorem 6.2, that the principal class is the identity element. The following examples show how to find the opposite (i.e. the inverse under composition) of a given class.

Example 6.6. If $D = b^2 - 4ac$ is even:

$$\left(ax^2 + bxy + cy^2\right)\left(az^2 - bzw + cw^2\right) = B^2 - \frac{D}{4}B'^2 = \left(axz - \frac{b}{2}xw + \frac{b}{2}yz - cyw\right)^2 -$$

$$-\frac{D}{4}\left(xw + yz\right)^2, \text{ with } \begin{pmatrix} B \\ B' \end{pmatrix} = \begin{pmatrix} a & -\frac{b}{2} & \frac{b}{2} & -c \\ 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix}$$

Example 6.7. If $D = b^2 - 4ac$ is odd:

$$\left(ax^2 + bxy + cy^2\right)\left(az^2 - bzw + cw^2\right) = B^2 + BB' + \frac{1-D}{4}B'^2 =$$

$$= \left(axz + \frac{-1-b}{2}xw + \frac{-1+b}{2}yz - cyw\right)^2 +$$

$$+ \left(axz + \frac{-1-b}{2}xw + \frac{-1+b}{2}yz - cyw\right)\left(xw + yz\right) + \frac{1-D}{4}\left(xw + yz\right)^2,$$

$$\text{with } \begin{pmatrix} B \\ B' \end{pmatrix} = \begin{pmatrix} a & \frac{-1-b}{2} & \frac{-1+b}{2} & -c \\ 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix}$$

Therefore, the opposite class of $ax^2 + bxy + cy^2$ is always the class of $ax^2 - bxy + cy^2$.

These computations that we have made are the proof of the following theorem:

Theorem 6.8. *The identity element of the class group is the principal class. The opposite of the class of a form $f(x, y) = ax^2 + bxy + cy^2$ is the class of the form $g(x, y) = ax^2 - bxy + cy^2$. The elements of order two are precisely the classes of ambiguous forms which are not the principal class.*

Proof. Assuming Theorem 6.2, the first two results follow from the above computations. If $f(x, y) = ax^2 + bxy + cy^2$ is an ambiguous form, it is improperly equivalent to itself, which means that it is properly equivalent to $g(x, y) = ax^2 - bxy + cy^2$. Therefore, the composition of the class of f with itself is the composition of the class of f with the class of g, which yields the class of the identity. Therefore, the class of f (which is supposed not to be the principal class), will have order two. The argument also works the other way round. □

Having seen some of the consequences of the main theorem, we proceed now to give a proof of it. But first, only one part of it, namely:

Proposition 6.9. *If h is a composition of f and g, and f', g' are properly equivalent to f and g, respectively, then h is also a composition of f' and g'.*

Proof. Because of the symmetric property of composition and the fact that any matrix in $\mathrm{SL}_2(\mathbb{Z})$ can be expressed as a finite product of $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $T^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ (S^{-1} is not needed since $S^{-1} = S^3$), it is sufficient to prove the result for $g' = g$ and f' being the result of the transformation by T^{-1}, T or S.

We will merge the first two cases into T^k, with $k = \pm 1$, but in fact the following works for any integer k.

Let's put some names to the coefficients:

$$\begin{gathered} f(x,y) = ax^2 + bxy + cy^2, \\ g(x,y) = a'x^2 + b'xy + c'y^2, \\ f'(x,y) = ax^2 + (b+2ka)xy + (ak^2 + bk + c)y^2, \\ h(B(x,y;z,w), B'(x,y;z,w)) = f(x,y)g(z,w) \\ \begin{pmatrix} B \\ B' \end{pmatrix} = \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix} \end{gathered}$$

Therefore, we have that $h\left(\hat{B}(x,y;z,w), \hat{B}'(x,y;z,w)\right) = f(x+ky,y)g(z,w) = f'(x,y)g(z,w)$, where

$$\begin{pmatrix} \hat{B} \\ \hat{B}' \end{pmatrix} = \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix} \begin{pmatrix} (x+ky)z \\ (x+ky)w \\ yz \\ yw \end{pmatrix} = \begin{pmatrix} p & q & kp+r & s \\ p' & q' & kp'+r' & s' \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix}$$

and the two extra conditions, which are $pq' - p'q = a$ and $p(kp' + r') - p'(kp + r) = a'$, are satisfied.

In the last case,

$$\begin{gathered} f(x,y) = ax^2 + bxy + cy^2, \\ g(x,y) = a'x^2 + b'xy + c'y^2, \\ f'(x,y) = cx^2 - bxy + ay^2, \\ h(B(x,y;z,w), B'(x,y;z,w)) = f(x,y)g(z,w) \\ \begin{pmatrix} B \\ B' \end{pmatrix} = \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix} \end{gathered}$$

Therefore, we have that $h\left(\hat{B}(x,y;z,w), \hat{B}'(x,y;z,w)\right) = f(-y,x)g(z,w) = f'(x,y)g(z,w)$,

where

$$\begin{pmatrix}\hat{B}\\ \hat{B}'\end{pmatrix}=\begin{pmatrix}p & q & r & s\\ p' & q' & r' & s'\end{pmatrix}\begin{pmatrix}-yz\\ -yw\\ xz\\ xw\end{pmatrix}=\begin{pmatrix}r & s & -p & -q\\ r' & s' & -p' & -q'\end{pmatrix}\begin{pmatrix}xz\\ xw\\ yz\\ yw\end{pmatrix}$$

and the two extra conditions, which are $rs'-r's=c$ and $-rp'+r'p=a'$, are satisfied, thanks to (6.2). □

We will now concern about the existence of a composition. Finding a composition of two arbitrary forms is a bit complicated. However, it is easy in some special cases:

Definition 6.10. Two primitive forms $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ of the same discriminant are said to be *close* if $\gcd\left(a,a',\frac{b+b'}{2}\right)=1$. They are said to be *joined* if $\gcd(a,a')=1$ and they are said to be *united* if $b=b'$ and $\gcd(a,a')=1$.

Note that all united forms are joined and all joined forms are close. Depending on the author, the term 'united' can also designate 'joined' or 'close' forms. I prefer using different names for distinguishing the three notions.

The concept of united forms is due to Dirichlet, who related Gauß's composition

If $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ are united, we have that their discriminant $D=b^2-4ac=b'^2-4a'c'$, so we have $c=ta'$ and $c'=ta$, for some integer t. Then the following identity shows us how to compute a composition:

Example 6.11.

$$\left(ax^2+bxy+cy^2\right)\left(a'z^2+b'zw+c'w^2\right)=\left(ax^2+bxy+ta'y^2\right)\left(a'z^2+bzw+taw^2\right)=$$

$$=aa'B^2+bBB'+tB'^2=$$

$$=aa'\left(xz-tyw\right)^2+b\left(xz-tyw\right)\left(axw+a'yz+byw\right)+t\left(axw+a'yz+byw\right)^2,$$

$$\text{with }\begin{pmatrix}B\\ B'\end{pmatrix}=\begin{pmatrix}1 & 0 & 0 & -t\\ 0 & a & a' & b\end{pmatrix}\begin{pmatrix}xz\\ xw\\ yz\\ yw\end{pmatrix}$$

Lemma 6.12. *If two forms $f(x,y)=ax^2+bxy+cy^2$ and $g(x,y)=a'x^2+b'xy+c'y^2$ of discriminant D are close, then the set of congruences*

$$\begin{cases}\lambda\equiv b & \mod 2a\\ \lambda\equiv b' & \mod 2a'\end{cases}$$

has a unique solution modulo $2\frac{aa'}{\gcd(a,a')}$.

Proof. Since $D = b^2 - 4ac = b'^2 - 4a'c'$, we have that $\left(\frac{b+b'}{2}\right)\left(\frac{b-b'}{2}\right) = ac - a'c'$. Now, $\gcd(a, a')$ clearly divides the right hand side, but it is relatively prime with $\frac{b+b'}{2}$. Therefore, it must divide the other factor and we can write $\frac{b-b'}{2} = -ak + a'l$ for some integers k, l. Rearranging that, we obtain that $\lambda := b + 2ak = b' + 2a'l$ is a solution, and so is $\lambda_z = b + 2ak + 2\frac{aa'}{\gcd(a,a')}z = b' + 2a'l + 2\frac{aa'}{\gcd(a,a')}z$ for all $z \in \mathbb{Z}$. Conversely, each pair of solutions λ, λ' satisfy that $\lambda - \lambda'$ is a multiple of $2a$ and $2a'$, so it is also a multiple of $\operatorname{lcm}(2a, 2a') = 2\frac{aa'}{\gcd(a,a')}$, so the solution is unique using this modulus. □

If we now apply the proper transformations $(x, y) \mapsto (x + ky, y)$ and $(x, y) \mapsto (x + ly, y)$, where k, l are the integers defined during the proof of the former lemma, f and g, respectively, get transformed into some forms $F(x, y) = ax^2 + \lambda xy + \mu y^2$, $G(x, y) = a'x^2 + \lambda xy + \mu' y^2$. If, f and g where also joined, we have that F and G are united.

Given two arbitrary forms f and g, if we could find some other forms f' and g' which are joined and such that f and f' are properly equivalent, and so are g and g', we would found a composition of the united forms F and G constructed from f' and g', and it would also be a composition of f and g. We will present such procedure shortly. This method of finding a composition of f and g is called Dirichlet composition, and it is the most common method of finding a composition of two given forms.

But first, we need the following lemma:

Lemma 6.13. *Let $f(x, y) = ax^2 + bxy + cy^2$ be a primitive form and let a' be an integer, Then, f primitively represents a number relatively prime to a'.*

Proof. For each prime p dividing a', find the pair $(\alpha_p, \gamma_p) \in \{(1, 0), (0, 1), (1, 1)\}$ such that $f(\alpha_p, \gamma_p)$ is relatively prime to p. This can be done because the fact that f is primitive implies that p cannot divide $f(1, 0) = a, f(0, 1) = c$ and $f(1, 1) = a + b + c$ at the same time. Then, find integers α and γ such that $\alpha \equiv \alpha_p$ and $\gamma \equiv \gamma_p$ for all p dividing a', which can be found using the Chinese remainder theorem. By construction, $f(\alpha, \gamma)$ will be relatively prime to a' since $f(\alpha, \gamma) \equiv f(\alpha_p, \gamma_p) \mod p$ for all $p \mid a'$. If α and γ are not relatively prime, divide both of them by their common factor, obtaining α' and γ', respectively. $f(\alpha', \gamma')$ thus obtained is a divisor of $f(\alpha, \gamma)$ and so it still will be relatively prime to a'. □

Using the lemma, one can find f' and g' equivalent to $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = a'x^2 + b'xy + c'y^2$, respectively, such that f' and g' are joined: Choose relatively prime α, γ such that $f(\alpha, \gamma)$ and a' have no common factors. Let β and δ be integers such that $\alpha\delta - \beta\gamma = 1$. Then $f'(x, y) = f(\alpha x + \beta y, \gamma x + \delta y)$ and $g'(x, y) = g(x, y)$ satisfy that $f'(1, 0) = f(\alpha, \gamma)$ and a' are relatively prime, so f' and g' are joined.

Proposition 6.14. *There always exists a composition between two primitive forms of the same discriminant.*

Proof. We first find, as shown above, two united forms F and G properly equivalent to f and g respectively. Then, we find a composition H of F and G. Proposition 6.9 ensures H is

also a composition of f and g. $\square$

We want to remark that the crucial point of this theorem is that there always exist united forms F and G properly equivalent to any given forms f and g of the same discriminant.

Let's now prove another part of Theorem 6.2:

Proposition 6.15. *Let $h(x,y)$ and $\hat{h}(x,y)$ be two compositions of two primitive forms $f(x,y) = = ax^2+bxy+cy^2$ and $g(x,y) = a'x^2+b'xy+c'y^2$. Then h and $\hat{h}$ are properly equivalent.*

Proof. Let $f(x,y)g(z,w) = h\,(B,B') = \hat{h}\left(\hat{B},\hat{B}'\right)$, where

$$\begin{pmatrix} B \\ B' \end{pmatrix} = \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix}; \qquad \begin{pmatrix} \hat{B} \\ \hat{B}' \end{pmatrix} = \begin{pmatrix} \hat{p} & \hat{q} & \hat{r} & \hat{s} \\ \hat{p}' & \hat{q}' & \hat{r}' & \hat{s}' \end{pmatrix} \begin{pmatrix} xz \\ xw \\ yz \\ yw \end{pmatrix}$$

and $pq'-p'q = \hat{p}\hat{q}'-\hat{p}'\hat{q} = a$, $pr'-p'r = \hat{p}\hat{r}'-\hat{p}'\hat{r} = a'$. Because of (6.2), we have that the other 2×2 determinants also coincide. Then, consider the following matrix:

$$\begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \\ \hat{p} & \hat{q} & \hat{r} & \hat{s} \\ \hat{p}' & \hat{q}' & \hat{r}' & \hat{s}' \end{pmatrix}$$

We claim that this matrix has rank two, which is achieved by some 2×2 determinant in the first two rows. Indeed, the 2×2 determinants formed by the first two rows, by (6.2), equal a, a', $\frac{b+b'}{2}$, $\frac{b-b'}{2}$, c' and c, which cannot be all zero. On the other hand, we have that the following 3×3 minor equals zero:

$$\begin{vmatrix} p & q & r \\ p' & q' & r' \\ \hat{p} & \hat{q} & \hat{r} \end{vmatrix} = \hat{p}(qr'-q'r) - \hat{q}(pr'-p'r) + \hat{r}(pq'-p'q) =$$

$$= \hat{p}(\hat{q}\hat{r}'-\hat{q}'\hat{r}) - \hat{q}(\hat{p}\hat{r}'-\hat{p}'\hat{r}) + \hat{r}(\hat{p}\hat{q}'-\hat{p}'\hat{q}) = \begin{vmatrix} \hat{p} & \hat{q} & \hat{r} \\ \hat{p}' & \hat{q}' & \hat{r}' \\ \hat{p} & \hat{q} & \hat{r} \end{vmatrix} = 0$$

The other fifteen 3×3 minors vanish analogously.

This means that there exist some rational numbers α, β, γ and δ such that

$$\begin{pmatrix} \hat{p} & \hat{q} & \hat{r} & \hat{s} \\ \hat{p}' & \hat{q}' & \hat{r}' & \hat{s}' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix} \text{ and so } \begin{pmatrix} \hat{B} \\ \hat{B}' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} B \\ B' \end{pmatrix}$$

which implies that $\hat{h}(\alpha x+\beta y, \gamma x+\delta y) = h(x,y)$.

We still need to prove that these α, β, γ and δ are integers and that $\alpha\delta - \beta\gamma = 1$. The latter condition can be seen by comparing the 2×2 minors in the former matrix equality. By looking at

$$\begin{pmatrix} \hat{p} & \hat{q} & \hat{r} & \hat{s} \\ \hat{p}' & \hat{q}' & \hat{r}' & \hat{s}' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} p & q & r & s \\ p' & q' & r' & s' \end{pmatrix}$$

we can focus on the first row and the first two columns:

$$\begin{cases} \hat{p} = \alpha p + \beta p' \\ \hat{q} = \alpha q + \beta q' \end{cases}$$

Applying Cramer's rule, we can solve for α and β and we see that their denominators have to divide $pq' - p'q = a$. Similarly, by looking at other pairs of columns, we see that these denominators must also divide a', $\frac{b+b'}{2}$, $\frac{b-b'}{2}$, c' and c, so they have to divide $\gcd\left(a, a', \frac{b+b'}{2}, \frac{b-b'}{2}, c', c\right) = \gcd\left(a, b, c, a', \frac{b+b'}{2}, c'\right) = 1$, so α and β are integers. Similarly, by observing the second row we see that γ and δ are also integers. This means that h and $\hat{h}$ are properly equivalent. $\square$

We are almost ready to prove Theorem 6.2, but first we need another lemma:

Lemma 6.16. *Given n primitive forms $f_1, f_2, \ldots, f_n$ of the same discriminant, we can find n forms $F_1, F_2, \ldots, F_n$ respectively equivalent such that they are pairwise united.*

Proof. Let $f_i(x, y) = a_i x^2 + b_i xy + c_i y^2$, for $i = 1, 2, \ldots, n$.

We will prove the result by induction on n. For $n = 1, 2$, it is true, so now suppose the lemma is true for $n = k$ and let's prove it for $n = k + 1$.

We apply the lemma to the first k forms, thus obtaining $F_i' = A_i x^2 + B'xy + C_i' y^2$, for $i = 1, 2, \ldots, k$, where all A_i's are pairwise relatively prime. Let A be the product of $A_1, A_2, \ldots, A_k$, and let $H' = F_k' \circ (\cdots \circ (F_3' \circ (F_2' \circ F_1')) \cdots)$, where $\circ$ denotes the composition of united forms found in Example 6.11 (it is well defined since the result of each composition is united to the other operand), so we have that necessarily $H'(x, y) = Ax^2 + B'xy + C'y^2$ for some C'.

Finally, let $F_{k+1}(x, y) = A_{k+1}x^2 + Bxy + C_{k+1}y^2$ and $H(x, y) = Ax^2 + Bxy + Cy^2$ be united and properly equivalent to f_k and H', respectively, found following Proposition 6.14. The transformation from H' to H is of the following form $(x, y) \mapsto (x + ly, y)$, so that $B = B' + 2lA$. Applying the transformation $(x, y) \mapsto \left(x + \frac{A}{A_i} ly, y\right)$ to $F_i'(x, y)$ we get $F_i(x, y) = A_i x^2 + Bxy + C_i y^2$.

Forms $F_1, F_2, \ldots, F_k, F_{k+1}$ are pairwise united (A_{k+1} is relatively prime to A and therefore to each of the other A_i's) and respectively properly equivalent to $f_1, f_2, \ldots, f_k, f_{k+1}$. $\square$

We are now able to present a proof to the main theorem of this section:

Proof.(Theorem 6.2)
We have already discussed the first point: A composition of two primitive forms of the same discriminant always exists due to Proposition 6.14. Different compositions are properly equivalent because of Proposition 6.15 and composition only depends on the proper class of the two operands thanks to Proposition 6.9.

Since we already know that the operation is symmetric, in order to prove it provides an abelian group structure we only need to prove associativity, the existence of an identity and the existence of inverses.

As for the associativity, let f, g and h be three primitive forms of the same discriminant. Let $F(x,y) = Ax^2+Bxy+Cy^2$ and let $G(x,y) = A'x^2+Bxy+Cy^2$ and $H(x,y) = A''x^2+Bxy+C''y^2$ be properly equivalent to f, g and h, respectively and be pairwise united. By making their discriminant equal, we see that $F(x,y) = Ax^2 + Bxy + tA'A''$, $G(x,y) = A'x^2 + Bxy + tAA''y^2$ and $H(x,y) = A''x^2 + Bxy + tAA'y^2$. We can now verify associativity:

If we first compose F and G, and then with H, we obtain $(AA')A''x^2 + Bxy + ty^2$, and if we first compose G and H, and then compose F with the result, we obtain $A(A'A'')x^2 + Bxy + ty^2$. Since the two results are properly equivalent (in fact equal), we have that composition is associative.

Example 6.4 and Example 6.5 convinced us that the principal class ought to be the identity by showing it is idempotent. Since we now are trying to see that composition induces a group, this is not enough, and we need to show that the principal class composed with any form leaves it invariant:

Let $f(x,y) = ax^2 + bxy + cy^2$ have even discriminant, and let $b = 2k$. Then, the form $i(x,y) = x^2 + 2kxy + \left(k^2 - \frac{D}{4}\right) y^2 = x^2 + bxy + acy^2$ is properly equivalent to the principal form via $(x,y) \mapsto (x+ky, y)$ and f and i are united. The composition is $ax^2 + bxy + cy^2 = f(x,y)$.

Let $f(x,y) = ax^2 + bxy + cy^2$ have odd discriminant, and let $b = 2k+1$. Then, the form $i(x,y) = x^2 + (2k+1)xy + \left(k^2 + k + \frac{1-D}{4}\right) y^2 = x^2 + bxy + acy^2$ is properly equivalent to the principal form via $(x,y) \mapsto (x+ky, y)$ and f and i are united. The composition is $ax^2 + bxy + cy^2 = f(x,y)$.

Since we already know that composing $ax^2 + bxy + cy^2$ with $ax^2 - bxy + cy^2$ yields the principal form, every element has an inverse. Therefore, we have proved the abelian group structure of the set of proper equivalence classes of primitive forms under composition. □

In regard of this group structure, we use the following terminology:

Definition 6.17. The group formed by the set of proper equivalence classes of primitive forms of discriminant D, with composition as operation, is called the *class group* of discriminant D if $D > 0$. If $D \leq 0$, the name *class group* is reserved to the subgroup of positive (semi)-definite

forms. We call $h(D)$, the *class number*, to the order of the class group.

Note that the class group is well-defined for negative discriminants since the composition of two positive definite forms is also positive definite because it represents products of numbers represented by the operands. The group of proper equivalence classes of primitive forms of discriminant $D \leq 0$, without asking the forms to be positive definite, is isomorphic to the product of the class group and $\{1, -1\} \cong \mathbb{Z}/2\mathbb{Z}$.

Theorem 6.18. *If $D = d^2 > 0$, the class group of discriminant D is isomorphic to $(\mathbb{Z}/d\mathbb{Z})^\times$. The class group of discriminant 0 is trivial.*

Proof. Suppose $d \neq 0$. It is more convenient to use right-diminished forms to prove this result: Each primitive proper class has a right-diminished representative $ax^2 + dxy$, with $\gcd(a, d) = 1$. Therefore, we can assign to each class the number $a \in (\mathbb{Z}/d\mathbb{Z})^\times$. Imagine we have two primitive right-diminished forms $f(x, y) = ax^2 + dxy$ and $g(x, y) = a'x^2 + dxy$. We can find some integer k such that $a' + kd$ is relatively prime to a (this can be done because $\gcd(a, d) = 1$). Then, $G(x, y) = (a' + kd)x^2 + dxy$ is united with f and it is properly equivalent to g by $(x, y) \mapsto (x, kx + y)$. The result of the composition is $a(a' + kd)x^2 + dxy$, which is properly equivalent to $a''x^2 + dxy$, being a'' the residue of $a(a' + kd)$, or aa', modulo d. This means that we assign to the composition the number $a'' \in (\mathbb{Z}/d\mathbb{Z})^\times$ which is the product of a and a' in that group. Therefore, the two groups are isomorphic.

If $d = 0$, then the only positive semi-definite right-diminished form is x^2, so the group has one element and it is trivial. □

Dirichlet, apart from providing a procedure to follow in order to compose forms, he gave another insight to composition: he related the group structure of classes of forms to a group of ideals of a certain order. We will study this connection in detail.

7 Ideal class group

As we anticipated before, we will establish a bijection between the group of proper classes and another group, in a way that composition of forms translates into the operation of the latter group. This group will be a group of classes of fractional ideals on an order, so we will proceed to define what an order is and what fractional ideals are, before imposing the group structure. In this section, the square roots of integer numbers will always be taken as positive or positive imaginary, unless stated otherwise.

Given a quadratic field $\mathbb{K} = \mathbb{Q}\left(\sqrt{D}\right)$, where D is not a square, we may consider its ring of algebraic integers $\mathcal{O}_{\mathbb{K}}$. First of all, let us characterize how this ring looks like.

Theorem 7.1. *Let d be the square-free integer such that $\mathbb{Q}\left(\sqrt{D}\right) = \mathbb{Q}\left(\sqrt{d}\right)$. Then the algebraic integers are precisely those of the form $x + y\tau_d$, where x and y are integers and*

$$\tau_d = \begin{cases} \sqrt{d} & \text{if } d \equiv 2, 3 \mod 4 \\ \frac{1+\sqrt{d}}{2} & \text{if } d \equiv 1 \mod 4 \end{cases}$$

Proof. First, we note that τ_d is always a root of $x^2 - d$ or $x^2 - x + \frac{1-d}{4}$, when $d \equiv 2, 3$ or $1 \mod 4$, respectively, and in each case the polynomial is integer and monic. Therefore, the numbers of the form $x + y\tau_d$ with integers x, y are always algebraic integers.

Conversely, let $\omega = a + b\sqrt{d}$, with rational numbers a, b, be an algebraic integer. Suppose it is an integer. Then, the statement holds trivially. Otherwise, its minimal polynomial is $(x - (a + b\sqrt{d}))(x - (a - b\sqrt{d})) = x^2 - 2ax + a^2 - db^2$, so $2a$ and $a^2 - db^2$ are integers. If a is an integer, then so is b, as d is square-free. Then either $[a] + [b]\sqrt{d}$ or $[a - b] + [2b]\frac{1+\sqrt{d}}{2}$ are representations of ω in the form required. On the other hand, if a is not an integer, let $a = \frac{A}{2}$, with A being an odd integer. Again, since d is square-free and b cannot be integer, it must happen that $b = \frac{B}{2}$, with B being an odd integer. Then $A^2 - dB^2$ must be a multiple of four, which forces $d \equiv 1 \mod 4$. We can write then $\omega = \left[\frac{A-B}{2}\right] + [B]\frac{1+\sqrt{d}}{2}$. □

Now we make two definitions: discriminant and order.

Definition 7.2.
We say an integer D is a *discriminant* if it is the discriminant of some quadratic form. Equivalently, D is a discriminant if $D \equiv 0, 1 \mod 4$.
We say that a discriminant D is *fundamental* if either D is odd and square-free or $D = 4m$ with $m \equiv 2, 3 \mod 4$ and square-free.

The term 'fundamental' comes from the fact that if a fundamental discriminant δ is written as a square times another discriminant: $\delta = \varphi^2 D$, then necessarily $D = \delta$ and $\varphi = 1$.

Note that every discriminant D can be written as $D = \varphi^2\delta$, where φ is a non-negative integer and δ is a fundamental discriminant. Moreover, this decomposition is unique.

Definition 7.3. An *order* in a number field $\mathbb{K}$ of degree n over $\mathbb{Q}$ is a subring $\mathcal{O}$ of $\mathbb{K}$ (and so contains 1) which is also a $\mathbb{Z}$-module of rank n.

Throughout this section, we will use angle brackets $\langle g_1, g_2, \ldots, g_n \rangle$ to designate the $\mathbb{Z}$-module generated by the $g_i's$.

Theorem 7.4. *An order $\mathcal{O}$ in $\mathbb{K}$ is always a subring of algebraic integers $\mathcal{O}_\mathbb{K}$.*

Proof. Consider $\alpha \in \mathcal{O}$. Since $\mathcal{O}$ is a ring, we have that $\mathbb{Z}[\alpha]$ is contained in $\mathcal{O}$. Since $\mathcal{O}$ is a finitely generated $\mathbb{Z}$-module, we have that $\mathbb{Z}[\alpha]$ is also so. The last statement is equivalent to α being an algebraic integer in $\mathbb{K}$. Therefore, $\mathcal{O} \subseteq \mathcal{O}_\mathbb{K}$. □

Conversely, $\mathcal{O}_\mathbb{K}$ is an order, called the *maximal order*. In the case $n = 2$, it is true since $\mathcal{O}_\mathbb{K} = \langle 1, \tau_d \rangle$. For greater n's the result is also true, but we do not prove it since we are only concerned about quadratic fields.

The above implies that, given any order $\mathcal{O}$ in $\mathbb{K}$, since both $\mathcal{O}$ and $\mathcal{O}_\mathbb{K}$ will be $\mathbb{Z}$-modules, or abelian groups, of rank n, we have that the index $[\mathcal{O}_\mathbb{K} : \mathcal{O}]$ as abelian groups is finite.

Definition 7.5. The *conductor* φ of an order $\mathcal{O}$ is the index $[\mathcal{O}_\mathbb{K} : \mathcal{O}]$.

Traditionally, the conductor is denoted by the letter f, but since we use that letter to designate quadratic forms, we may use the letter φ instead. Let's now turn our head into the quadratic case. If the same result can be applied to a general field, we will announce it.

Proposition 7.6. *If $\mathcal{O}$ in a quadratic field has conductor φ, then $\mathcal{O} = \langle 1, \varphi\tau_d \rangle$. In every quadratic field there exists a unique order of conductor φ.*

Proof. Since the quotient group $\mathcal{O}_\mathbb{K}/\mathcal{O}$ has (group) order φ, we have that each element of $\mathcal{O}_\mathbb{K}$ added φ times becomes the zero element, that is, lies in $\mathcal{O}$. Therefore, we have $\varphi\mathcal{O}_\mathbb{K} \subseteq \mathcal{O}$. In particular, $\varphi\tau_d \in \mathcal{O}$ and $\langle 1, \varphi\tau_d \rangle \subseteq \mathcal{O}$. Since $\langle 1, \varphi\tau_d \rangle$ has also index φ, both groups must be equal.

It is now clear that there is at most one order for each conductor φ, but we still need to prove that there exists one. That is, we need to prove that $\langle 1, \varphi\tau_d \rangle$ is always an order.

$\langle 1, \varphi\tau_d \rangle$ is obviously a $\mathbb{Z}$-module of rank 2, so we only need to see that it is a subring. The conditions of being an additive subgroup and containing 1 are obvious, and multiplication is closed since the fact that τ_d is an algebraic integer of degree two implies that $(\varphi\tau_d)^2 \in \varphi^2 \langle 1, \tau_d \rangle \subseteq \langle 1, \varphi\tau_d \rangle$.

Therefore, for each φ there exists a unique order of conductor φ in a given quadratic field. □

Definition 7.7. The *discriminant* of an order $\mathcal{O}$ in a quadratic field $\mathbb{K}$ is the quantity

$$\left(\det\begin{pmatrix}\mu & \nu\\ {}^{\sigma}\mu & {}^{\sigma}\nu\end{pmatrix}\right)^2,$$

where σ is the non-trivial automorphism of $\mathbb{K}$ and $\{\mu,\nu\}$ is a basis of $\mathcal{O}$ as a $\mathbb{Z}$-module.

In general, one can define the discriminant of an order in any field, in a similar fashion: in each column of the matrix, write the elements of the basis, and apply σ_i to row i, where $\sigma_1,\sigma_2,\ldots,\sigma_n$ are the embeddings of $\mathbb{K}$ in $\mathbb{C}$, which now need not be automorphisms, in any order. Once the matrix is completed, take the determinant and square it.

Proposition 7.8. *The discriminant of an order is well-defined.*

Proof. This is a consequence of the fact that every change of basis will have integer coefficients and be invertible. This means that

$$\begin{pmatrix}\mu'\\ \nu'\end{pmatrix}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}\mu\\ \nu\end{pmatrix},\begin{pmatrix}{}^{\sigma}\mu'\\ {}^{\sigma}\nu'\end{pmatrix}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}{}^{\sigma}\mu\\ {}^{\sigma}\nu\end{pmatrix}\implies\begin{pmatrix}\mu' & \nu'\\ {}^{\sigma}\mu' & {}^{\sigma}\nu'\end{pmatrix}=\begin{pmatrix}\mu & \nu\\ {}^{\sigma}\mu & {}^{\sigma}\nu\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{\top};$$

with $\alpha\delta-\beta\gamma=\pm1$. Therefore, the determinant in the definition can only change by a sign, which is then taken care of by the square. The same argument works for an order in a general field. □

Since we know a basis for the ring of algebraic integers and for any order in terms of its conductor, we may calculate their respective discriminants:

Proposition 7.9.
The discriminant of the ring of algebraic integers $\mathcal{O}_{\mathbb{Q}(\sqrt{d})}$, with d square-free is d if $d\equiv 1 \mod 4$ and $4d$ if $d\equiv 2,3 \mod 4$.
The discriminant of an order of index φ is equal to φ^2 times the discriminant of the ring of algebraic integers where it is contained.

Proof. The first statement comes from

$$\left(\det\begin{pmatrix}1 & \sqrt{d}\\ 1 & -\sqrt{d}\end{pmatrix}\right)^2=4d;\qquad\left(\det\begin{pmatrix}1 & \frac{1+\sqrt{d}}{2}\\ 1 & \frac{1-\sqrt{d}}{2}\end{pmatrix}\right)^2=d$$

The second statement is obvious since using the basis $\{1,\varphi\tau_d\}$, the factor φ can be pulled out of the determinant. □

Note that the discriminant of the ring of integers $\mathcal{O}_{\mathbb{K}}$ is always a fundamental discriminant. and the discriminant of any order is always a discriminant. Also,

Corollary 7.10. *For each discriminant D that is not a square, there exists a unique order of discriminant D.*

Proof. Let $D = \varphi^2\delta$, being δ fundamental. For existence, let d be δ or $\frac{\delta}{4}$, depending on the parity of δ, so $d \neq 1$ is square-free. It is sufficient to consider the order $\langle 1, \varphi\tau_d \rangle$ in $\mathbb{Q}\left(\sqrt{d}\right) = \mathbb{Q}\left(\sqrt{\delta}\right) = \mathbb{Q}\left(\sqrt{D}\right)$, which is the only order of conductor φ in that field. Any other order (including those over a different quadratic field) would lead to a different decomposition of $D = \delta\varphi^2$, which is impossible. □

Because of this result, we may call $\mathcal{O}_D$ the unique order of discriminant D.

We will now focus on the ideals of a given order. We are now getting closer to constructing the group.

Proposition 7.11. *The non-zero ideals of an order $\mathcal{O}$ have rank 2, and they have finite index as additive subgroups therefore.*

Proof. The second result follows from the first one.
Let I be a non-zero ideal of $\mathcal{O} = \langle 1, \varphi\tau_d \rangle$. Since I is a non-trivial subgroup of a free abelian group of rank 2, it can only have rank 1 or 2. Suppose that it has rank 1, that is, $I = \mathbb{Z} \cdot \mu$. Then, since I is an ideal of $\mathcal{O}$ and $\mu \in I$, we have that $\varphi\tau_d\mu \in I = \mathbb{Z} \cdot \mu$, which implies the contradictory result that $\varphi\tau_d$ is an integer. Therefore, I has rank 2. □

Definition 7.12. A *fractional ideal* of $\mathcal{O}$ is a set of the form ωI where $\omega \in \mathbb{K}^\times$ and I is an non-zero ideal of $\mathcal{O}$.

Note that any ideal is also a fractional ideal.

Proposition 7.13. *Given a fractional ideal $\mathfrak{a}$ of $\mathcal{O}$, there exists an element $\mu \in \mathcal{O}$ such that $\mu \neq 0$ and $\mu\mathfrak{a}$ is an ideal of $\mathcal{O}$.*

Proof. We could even force $\mu \in \mathbb{Z}$ and the statement would still be true.

Let $\mathcal{O} = \langle 1, \varphi\tau_d \rangle$. Since the set of algebraic integers is $\mathcal{O}_\mathbb{K} = \langle 1, \tau_d \rangle$, each element $\omega \in \mathbb{K}$ can be written as a quotient of two elements of $\mathcal{O}_\mathbb{K}$. Eliminating the square roots in the denominator, we can write $\omega = \alpha\tau_d + \beta$, where α, β are rational numbers. Let $0 \neq n \in \mathbb{Z}$ be such that $n\alpha$ and $n\beta$ are integers. Then $n\varphi\omega \in \langle \varphi, \varphi\tau_d \rangle \subseteq \langle 1, \varphi\tau_d \rangle = \mathcal{O}$. Therefore, if $\mathfrak{a} = \omega I$, taking $\mu = n\varphi$ suffices. □

Now we get a sense of why they are called fractional ideals: multiplying a fraction by the appropriate number yields an integer, while multiplying the whole fractional ideal by an appropriate integer yields an ideal.

Definition 7.14. Given a fractional ideal $\mathfrak{a}$ of $\mathcal{O}$, we call the set $\{\omega \in \mathbb{K} \ : \ \omega\mathfrak{a} \subseteq \mathfrak{a}\}$, the ring of multipliers of $\mathfrak{a}$.

Proposition 7.15. *The ring of multipliers of a given fractional ideal is indeed a ring.*

Proof. The proof follows from the three following observations: $1\mathfrak{a} \subseteq \mathfrak{a}$, $(\omega \pm \omega')\mathfrak{a} \subseteq \subseteq \omega\mathfrak{a} \pm \omega'\mathfrak{a} \subseteq \mathfrak{a} + \mathfrak{a} \subseteq \mathfrak{a}$ and $\omega\omega'\mathfrak{a} \subseteq \omega\mathfrak{a} \subseteq \mathfrak{a}$. □

Definition 7.16.
A fractional ideal $\mathfrak{a}$ is said to be *proper* if its ring of multipliers is $\mathcal{O}$.
A fractional ideal $\mathfrak{a}$ is said to be *invertible* if there exists another fractional ideal $\mathfrak{b}$ such that $\mathfrak{ab} = \mathcal{O}$.

Before going on, notice that we always have $\mathcal{O} \subseteq \{\xi \in \mathbb{K} \ : \ \xi\mathfrak{a} \subseteq \mathfrak{a}\}$ since if $\mu \in \mathcal{O}$, then $\mu\mathfrak{a} = \mu\omega I \subseteq \omega I = \mathfrak{a}$, so the definition of 'proper' only asks for the other inclusion. Proper ideals let us recover the order they belong to by computing its ring of multipliers.

Lemma 7.17. *Let ω be a solution of $ax^2 + bx + c = 0$, where a, b and c are integers and $\gcd(a, b, c) = 1$. Then, $\mathfrak{a} = \langle 1, \omega\rangle$ is a proper fractional ideal of the order $\mathcal{O} = \langle 1, a\omega\rangle$, and so is $\psi \langle 1, \omega\rangle$ for any $\psi \in \mathbb{K}^\times$.*

Proof. First, we note that since $a\omega$ is an algebraic integer, $\mathcal{O}$ is an order, as expected. Since $\langle a, a\omega\rangle$ is an ideal of $\mathcal{O}$ (which can be proved using $a^2\omega^2 = -ab\omega - ac \in \langle a, a\omega\rangle$), we have that $\psi \langle 1, \omega\rangle$ is a fractional ideal. Now we need to show it is proper. Consider the ring of multipliers $\{\chi \in \mathbb{K} \ : \ \chi\psi \langle 1, \omega\rangle \subseteq \psi \langle 1, \omega\rangle\} = \{\chi \in \mathbb{K} \ : \ \chi \langle 1, \omega\rangle \subseteq \langle 1, \omega\rangle\}$. Each χ in this set can be written as $\chi \cdot 1 = m\omega + n$, so

$$\chi\omega = m\omega^2 + n\omega = m\left(-\frac{b}{a}\omega - \frac{c}{a}\right) + n\omega = -m\frac{c}{a} + \left(-m\frac{b}{a} + n\right)\omega \in \langle 1, \omega\rangle$$

The only way the condition can be met is if $a \mid m$, since a cannot have common factors with both b and c. Therefore, the desired set is

$$\{\chi \in \mathbb{K} \ : \ \chi \langle 1, \omega\rangle \subseteq \langle 1, \omega\rangle\} = \{m\omega + n \ : \ a \mid m\} = \langle 1, a\omega\rangle = \mathcal{O}$$

and so the fractional ideals are proper. □

Proposition 7.18. *Proper fractional ideals are precisely those that are invertible.*

Proof. Let $\mathfrak{a}$ be an invertible fractional ideal (with inverse $\mathfrak{b}$) and let ω in $\mathbb{K}$ satisfy $\omega\mathfrak{a} \subseteq \mathfrak{a}$. Then, we have $\omega\mathcal{O} = \omega\mathfrak{ab} \subseteq \mathfrak{ab} = \mathcal{O}$, which implies $\omega \cdot 1 = \omega \in \mathcal{O}$. Therefore, $\mathfrak{a}$ is proper.

Conversely, since we know that every non-zero ideal has rank 2 as an abelian group, so does every fractional ideal. Let $\mathfrak{a} = \langle \psi, \xi\rangle$ be a proper fractional ideal of $\mathcal{O}$ and define $\omega = \frac{\xi}{\psi}$, so that $\mathfrak{a} = \psi \langle 1, \omega\rangle$. ξ and ψ are incommensurable since $\mathfrak{a}$ has rank 2. Let $ax^2 + bx + c \in \mathbb{Z}[x]$, with $\gcd(a, b, c) = 1$ be the minimal polynomial of ω. Then $\mathfrak{a} = \psi \langle 1, \omega\rangle$ is a proper fractional ideal of both $\mathcal{O}$ and $\langle 1, a\omega\rangle$, which means

$$\mathcal{O} = \{\chi \in \mathbb{K} \ : \chi\psi \langle 1, \omega\rangle \subseteq \psi \langle 1, \omega\rangle\} = \langle 1, a\omega\rangle$$

Since the other root ${}^{\sigma}\omega$ satisfies $a\omega + a\,{}^{\sigma}\omega = -b \in \mathbb{Z}$, we see that $\mathcal{O} = \langle 1, a\omega\rangle = \langle 1, a\,{}^{\sigma}\omega\rangle$, so the fractional ideal ${}^{\sigma}\mathfrak{a} = {}^{\sigma}\psi\,\langle 1,\, {}^{\sigma}\omega\rangle$ of $\mathcal{O}$ is also proper.

The fractional ideal $\mathfrak{b} = \frac{a}{\psi\,{}^{\sigma}\psi}\,{}^{\sigma}\mathfrak{a}$ is the inverse we are looking for:

$$\mathfrak{a}\mathfrak{b} = a\,\langle 1, \omega\rangle \cdot \langle 1,\, {}^{\sigma}\omega\rangle = \langle a, a\omega, a\,{}^{\sigma}\omega, a\omega\,{}^{\sigma}\omega\rangle =$$

$$= \langle a, a\omega, a\omega + a\,{}^{\sigma}\omega, a\omega\,{}^{\sigma}\omega\rangle = \langle a, a\omega, -b, c\rangle = \langle 1, a\omega\rangle = \mathcal{O}$$

□

Now, we are able to define a group structure on the set of proper fractional ideals:

Theorem 7.19. *The set of non-zero proper fractional ideals has an abelian group structure, with multiplication of fractional ideals as its operation. The identity is the fractional ideal $\mathcal{O} = (1)$, which is principal.*

Proof. Commutativity and associativity of ideal multiplication is immediate. Since $\mathfrak{a}\mathcal{O} = \mathfrak{a}$ for all fractional ideals $\mathfrak{a}$, we have that $\mathcal{O}$ is the identity element. The existence of inverses has been proven in the last proposition, so the only thing left to see is that given two proper fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$, the product is also proper. But this is true because the product of invertible elements is invertible, and so proper. □

Definition 7.20.
The *norm* $N(I)$ of an non-zero ideal I of an order $\mathcal{O}$ is the index $[\mathcal{O} : I] \in \mathbb{Z}^+$.
The *norm* $N(\omega)$ of an element $\omega \in \mathbb{K}$ is $\omega\,{}^{\sigma}\omega \in \mathbb{Q}$.
The *norm* $N(\mathfrak{a})$ of a non-zero proper fractional ideal $\mathfrak{a} = \omega I$ is defined as $|N(\omega)|N(I) \in \mathbb{Q}^+$.

Proposition 7.21. *The following statements are true:*

1. *The norm of proper ideals is multiplicative, and so is the norm of elements.*
2. *$N((\omega)) = |N(\omega)|$ for all ω in $\mathcal{O}$.*
3. *The norm of a proper fractional ideal is well-defined and it is multiplicative.*

Proof. The fact that the norm of elements is multiplicative is obvious.

Let $\mathcal{O} = \langle 1, u\rangle$ and let $u^2 + Bu + C = 0$. First of all, let's prove 2.:

Let $\omega = \gamma u + \delta$. Since $\omega \in \mathcal{O}$, we have $\omega u = -\gamma(Bu + C) + \delta u = (\delta - B\gamma)u + (-C\gamma) =: \alpha u + \beta$. Then $(\omega) = \langle \omega, u\omega\rangle = \langle \gamma u + \delta, \alpha u + \beta\rangle$. It is a well-known result in the theory of lattices that $N((\omega)) = [\mathcal{O} : (\omega)] = [\langle 1, u\rangle : \langle \gamma u + \delta, \alpha u + \beta\rangle] = |\alpha\delta - \beta\gamma|$. On the other hand

$$N(\omega) = (\gamma u + \delta)\,(\gamma\,{}^{\sigma}u + \delta) = \gamma^2 u\,{}^{\sigma}u + \gamma\delta\,(u + {}^{\sigma}u) + \delta^2 = \gamma^2 C - B\gamma\delta + \delta^2 = \alpha\delta - \beta\gamma$$

Let I be an ideal of $\mathcal{O}$ and let $\mu \in \mathcal{O}$. Since we have $\mu I \subseteq (\mu) \subseteq \mathcal{O}$, we have that $[\mathcal{O} : (\mu)][(\mu) : \mu I] = [\mathcal{O} : \mu I]$. Since the quotient $\mathcal{O}/I$ is isomorphic to $\mu\mathcal{O}/\mu I = (\mu)/\mu I$, we have that $[(\mu) : \mu I] = |(\mu)/\mu I| = |\mathcal{O}/I| = [\mathcal{O} : I]$. Therefore, we have that $N((\mu))N(I) = N(\mu I)$.

Now, let $I = \langle \psi, \xi \rangle$ for some $\psi, \xi \in \mathcal{O}$ be a proper ideal. Let $\omega = \frac{\xi}{\psi}$ and let $ax^2 + bx + c$ be the minimal polynomial of ω. Then $I = \psi \langle 1, \omega \rangle$ being proper and Lemma 7.17 imply that $\mathcal{O} = \langle 1, a\omega \rangle$. Then, the ideal $aI = \psi \langle a, a\omega \rangle$ and

$$a^2 N(I) = N((a))N(I) = N(aI) = N((\psi))N(\langle a, a\omega \rangle) = aN((\psi)),$$

where we have used $N((a)) = |N(a)| = |a\,{}^{\sigma}a| = a^2$. Therefore, $N(I) = \frac{|N(\psi)|}{a}$. On the other hand, we have that $I\,{}^{\sigma}I = \frac{\psi\,{}^{\sigma}\psi}{a} \langle a, a\omega, a\,{}^{\sigma}\omega, a\omega\,{}^{\sigma}\omega \rangle = \frac{N(\psi)}{a} \langle 1, a\omega \rangle = N(I)\mathcal{O}$, and

$$N(IJ)\mathcal{O} = IJ\,{}^{\sigma}(IJ) = I\,{}^{\sigma}I \cdot J\,{}^{\sigma}J = N(I)\mathcal{O} \cdot N(J)\mathcal{O} = N(I)N(J)\mathcal{O},$$

for any proper ideals I and J. Since both $N(IJ)$ and $N(I)N(J)$ are positive integers, it must be that $N(I)N(J) = N(IJ)$. We have thus proved 1.

Suppose $\mathfrak{a} = \omega I = \psi J$. Let $n, m \in \mathcal{O}$ be non-zero numbers such that $n\omega$ and $m\psi$ belong to $\mathcal{O}$. Then

$$|N(m)N(n)||N(\omega)|N(I) = |N(n\omega)N(m)|N(I) = N(nm\omega I) =$$
$$= N(nm\psi J) = |N(m\psi)N(n)|N(J) = |N(m)N(n)||N(\psi)|N(J)$$

This implies that $N(\mathfrak{a})$ is well defined, and the multiplicative property follows from the same property on the other two norms, so 3. holds. □

One can see that the principal fractional ideals (that is, those of the form $\omega \cdot (1)$) form a subgroup. Moreover, the principal fractional ideals with a generator of positive norm (that is, those that can be written as $\omega \cdot (1)$ with $N(\omega) > 0$) also form a subgroup.

Note that in imaginary quadratic fields, the norm is the square of the complex absolute value, which is always positive, so it is necessary to be in a real quadratic field to have negative norm. Also observe that a principal fractional ideal with a generator of positive norm may also be generated by an element of negative norm. That is why we used the terms ≪can be written as≫.

Definition 7.22. We call *ideal class group* the quotient group of fractional ideals modulo principal fractional ideals. We say that two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ are *equivalent* (and we write $\mathfrak{a} \sim \mathfrak{b}$) if they reduce to the same element in the quotient.
We call *narrow ideal class group* the quotient group of fractional ideals modulo principal fractional ideals with a generator of positive norm. We say that two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ are *narrowly equivalent* if they reduce to the same element in the quotient.

Because of the inclusion relation between the respective subgroups, we have that narrowly equivalent fractional ideals are also equivalent. Also, the difference between the two only exists if there is a fractional principal ideal that any generator of it has negative norm:

Theorem 7.23. *If an order is contained in an imaginary quadratic field, the narrow ideal class group and the ideal class group coincide. If the order is contained in a real quadratic field, the narrow ideal class group and the ideal class group coincide precisely when there is a unit of norm* -1.

Proof. The first part of the statement holds because for all $\omega = \alpha + \beta\sqrt{d} \neq 0$ in a quadratic field $\mathbb{Q}\left(\sqrt{d}\right)$ with $d < 0$, we have $N(\omega) = \alpha^2 - d\beta^2 > 0$.

Let $\mathbb{K} = \mathbb{Q}\left(\sqrt{d}\right)$, be a real quadratic field ($d > 0$). Since the norm in $\mathcal{O}_{\mathbb{K}}$ takes integer values, is multiplicative and satisfies $N(1) = 1$, the norm of a unit can only be ± 1. If ε is a unit of norm -1, then $-N(\omega) = N(\omega\varepsilon)$ implies that every principal ideal $(\omega) = (\omega\varepsilon)$ admits a generator of positive norm, which means that both notions agree.

Finally, if no such unit exists, then the ideal $\left(\varphi\sqrt{d}\right)$ of the order of conductor φ is not generated by any element of positive norm. □

Proposition 7.24. *Given integers* a, b, c *with* $\gcd(a, b, c) = 1$ *and* $D = b^2 - 4ac$ *not a square, the set* $\left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle$ *is a proper ideal of* $\mathcal{O}_D$.

Proof. Let $\omega = \frac{-b+\sqrt{D}}{2a}$. Then, we know that $a\langle 1, \omega\rangle = \langle a, a\omega\rangle = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle$ is a proper ideal of $\langle 1, a\omega\rangle$. We only need to show that $\langle 1, a\omega\rangle = \left\langle 1, \frac{-b+\sqrt{D}}{2} \right\rangle$ is the same as $\mathcal{O}_D$. Since there is only one order of discriminant D, the result follows from the next computation:

$$\left(\det \begin{pmatrix} 1 & a\omega \\ 1 & a\,{}^{\sigma}\omega \end{pmatrix}\right)^2 = \left(\frac{-b-\sqrt{D}}{2} - \frac{-b+\sqrt{D}}{2}\right)^2 = D$$

□

Definition 7.25. Given a positive definite/indefinite primitive form $f(x, y) = ax^2 + bxy + cy^2$ of discriminant D, we say that the ideal $I_f = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle \mu$ of $\mathcal{O}_D$ is an *associated ideal of* f if $\mu \in \mathcal{O}$ and $N(\mu)$ has the same sign as a.

It is trivial to see that there exists at least one associated ideal to such forms and that all associated ideals of a given form are narrowly equivalent. This definition will then be extended to classes, that is, we will see that properly equivalent forms are associated with narrowly equivalent ideals and that the association between classes will be one to one.

Proposition 7.26. *If two positive definite/indefinite primitive forms* $f(x, y) = ax^2 + bxy + cy^2$ *and* $g(x, y) = a'x^2 + b'xy + c'y^2$ *are properly equivalent, their associated ideals are narrowly equivalent.*

Proof. Since $\mathrm{SL}_2(\mathbb{Z})$ is generated by two matrices S and T, it is sufficient to prove the result when g is the transformation of f under the action of these matrices: $g(x,y) = cx^2 - bxy + ay^2$ and $g(x,y) = ax^2 + (b+2a)xy + (c+b+a)y^2$.

Let D be their determinant. For the first case, let μ and μ' be numbers whose norm has the same sign as a and c, respectively. We have that

$$I_f = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle \mu \sim \left\langle a\frac{b+\sqrt{D}}{2}, \frac{-b+\sqrt{D}}{2} \cdot \frac{b+\sqrt{D}}{2} \right\rangle \mu' = \left\langle a\frac{b+\sqrt{D}}{2}, \frac{D-b^2}{4} \right\rangle \mu' =$$

$$= \left\langle a\frac{b+\sqrt{D}}{2}, -ac \right\rangle \mu' = \left\langle ac, a\frac{b+\sqrt{D}}{2} \right\rangle \mu' \sim \left\langle c, \frac{b+\sqrt{D}}{2} \right\rangle \mu' = I_g$$

Since $N(a) = a^2 > 0$ and $N\left(\frac{b+\sqrt{D}}{2} \cdot \frac{\mu'}{\mu}\right) = \frac{b^2-D}{4}\frac{N(\mu')}{N(\mu)} = ac\frac{N(\mu')}{N(\mu)} > 0$, all equivalences presented were also narrow.

For the second case, we have

$$I_f = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle \mu = \left\langle a, -a + \frac{b+\sqrt{D}}{2} \right\rangle \mu = \left\langle a, \frac{-(b+2a)+\sqrt{D}}{2} \right\rangle \mu = I_g$$

□

Given an ideal, we would want to find a form who has that ideal as an associate. This can always be done:

Proposition 7.27. *The proper fractional ideal $\langle \psi, \xi \rangle$ of an order $\mathcal{O}$ in an imaginary (resp. real) quadratic field is narrowly equivalent to some ideal I_f, for some positive definite (resp. indefinite) primitive form f.*
In any case, the discriminant of the order and the form coincide.

Proof. Let $\omega = \frac{\xi}{\psi}$. Let $ax^2 + bx + c$, with $\gcd(a,b,c) = 1$ and $a > 0$, be the minimal polynomial of ω and therefore, $\omega = \frac{-b\pm\sqrt{D}}{2a}$, being $D = b^2 - 4ac$. We also have that $\mathcal{O} = \langle 1, a\omega \rangle$. Let $f(x,y) = ax^2 + bxy + cy^2$, $g(x,y) = ax^2 - bxy + cy^2$, $h(x,y) = -ax^2 - bxy - cy^2$ and $j(x,y) = -ax^2 + bxy - cy^2$.

Then, depending on the sign of the square root in $\omega = \frac{-b\pm\sqrt{D}}{2a}$, either

$$\langle \psi, \xi \rangle = \psi \langle 1, \omega \rangle \sim \langle 1, \omega \rangle \sim a \langle 1, \omega \rangle = a \left\langle 1, \frac{-b+\sqrt{D}}{2a} \right\rangle = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle \sim I_f \sim I_j$$

or

$$\langle \psi, \xi \rangle = \psi \langle 1, \omega \rangle \sim \langle 1, \omega \rangle \sim -a \langle 1, \omega \rangle = -a \left\langle 1, \frac{-b-\sqrt{D}}{2a} \right\rangle = \left\langle a, \frac{b+\sqrt{D}}{2} \right\rangle \sim I_g \sim I_h$$

If we are in an imaginary quadratic field, then ω must be non-real and f and g will be positive definite. Since, if that is the case, narrow equivalence and equivalence are the same, we are done. Otherwise, ω is real and all four forms are indefinite. Choose between f and j or between g and h, depending on the sign of $N(\psi)$.

The discriminant of $\mathcal{O} = \langle 1, a\omega \rangle = \left\langle 1, \frac{-b \pm \sqrt{D}}{2} \right\rangle$ is D, which is the discriminant of all four forms. $\square$

What we have done so far tells us that the mapping from primitive non-degenerate forms of discriminant D to ideals of $\mathcal{O}_D$ can be lifted to the level of classes. The last proposition states that the mapping is surjective. We still need to see injectivity:

Theorem 7.28. *The association of positive definite/indefinite primitive forms of discriminant D and ideals of the order of discriminant D induces a bijection between the proper classes and the narrow classes.*

Proof. As we just said, we have already proven that the mapping can be done at the level of classes, and we also proved that it is surjective. Our only concern now is injectivity:

Let $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = a'x^2 + b'xy + c'y^2$ be positive definite/indefinite primitive forms of discriminant D. Let μ and μ' have norms whose signs are the same as a and a', respectively. Suppose that I_f is narrowly equivalent to I_g. Then, since $a, a' \in \mathbb{Z}$ have positive norm, we have that $\left\langle 1, \frac{-b+\sqrt{D}}{2a} \right\rangle \mu$ and $\left\langle 1, \frac{-b'+\sqrt{D}}{2a'} \right\rangle \mu'$ are narrowly equivalent, so there exists some $\lambda \in \mathbb{Q}\left(\sqrt{D}\right)$, with $N(\lambda)$ having the same sign as $N\left(\frac{\mu'}{\mu}\right)$ (or aa'), such that $\langle 1, \omega \rangle = \lambda \langle 1, \omega' \rangle$, being $\omega = \frac{-b+\sqrt{D}}{2a}$ and $\omega' = \frac{-b'+\sqrt{D}}{2a'}$.

So, there exist some integers $\alpha, \beta, \gamma, \delta$ with $\alpha\delta - \beta\gamma = \pm 1$ such that $\lambda\omega' = \alpha\omega + \beta$ and $\lambda = \gamma\omega + \delta$. Therefore, we have that $\omega' = \frac{\alpha\omega+\beta}{\gamma\omega+\delta}$, so

$$0 = g(\omega', 1) = g\left(\frac{\alpha\omega+\beta}{\gamma\omega+\delta}, 1\right) = \left(\frac{1}{\gamma\omega+\delta}\right)^2 g(\alpha\omega+\beta, \gamma\omega+\delta)$$

This means that $h(x,y) = g(\alpha x + \beta y, \gamma x + \delta y)$ is primitive and satisfies $h(\omega, 1) = 0$, so $h(x,1) = \pm f(x,1)$ is the minimal polynomial of ω, and so $\pm f(x,y) = h(x,y) = g(\alpha x + \beta y, \gamma x + \delta y)$.

On the other hand,

$$\omega' = \frac{\alpha\omega+\beta}{\gamma\omega+\delta} = \frac{-\alpha b + \alpha\sqrt{D} + 2a\beta}{-\gamma b + \gamma\sqrt{D} + 2a\delta} = \frac{\left((2a\beta - b\alpha) + \alpha\sqrt{D}\right)\left((2a\delta - b\gamma) - \gamma\sqrt{D}\right)}{\left((2a\delta - b\gamma) + \gamma\sqrt{D}\right)\left((2a\delta - b\gamma) - \gamma\sqrt{D}\right)} =$$

$$= \frac{((2a\beta - b\alpha)(2a\delta - b\gamma) - \alpha\gamma D) + (2a\alpha\delta - b\alpha\gamma + b\alpha\gamma - 2a\beta\gamma)\sqrt{D}}{(2a\delta - b\gamma)^2 - \gamma^2 D} =$$

$$= q + \frac{2a(\alpha\delta - \beta\gamma)\sqrt{D}}{4a^2\delta^2 - 4ab\gamma\delta + b^2\gamma^2 - b^2\gamma^2 + 4ac\gamma^2} = q + \frac{2a(\alpha\delta - \beta\gamma)\sqrt{D}}{4a\left(a\delta^2 - b\gamma\delta + c\gamma^2\right)} = q + \frac{\alpha\delta - \beta\gamma}{2f(\delta, -\gamma)}\sqrt{D}$$

for some rational q.

We know that $f(\delta,-\gamma)=\pm g(\alpha\delta-\beta\gamma,0)=\pm a'$. Since aa' and the denominator $\pm 4aa'=$ $=4af(\delta,-\gamma)=4a\left(a\delta^2-b\gamma\delta+c\gamma^2\right)=N(\gamma\omega+\delta)=N(\lambda)$ have the same sign, we see that the sign we have to take is the positive one. However, we also know that $\omega'=\frac{-b'+\sqrt{D}}{2a'}$, which implies $q=-\frac{b'}{2a'}$ and $\frac{\alpha\delta-\beta\gamma}{2f(\delta,-\gamma)}=\frac{1}{2a'}$. This means that $\alpha\delta-\beta\gamma=1$ and so f and g are properly equivalent.

Therefore, the mapping at the level of classes is injective. □

We have now established a bijection between two sets, both of which had an additional structure of abelian group. Our goal now is to prove that the operations agree under this bijection.

Theorem 7.29. *The bijections presented are, in fact, abelian group isomorphisms between the class group of discriminant D, under composition, and the narrow ideal class group of discriminant D, under ideal multiplication, for each non-square discriminant D.*

Proof. Suppose the composition of the proper classes of primitive forms f and g is the class of h, all forms having discriminant D. We need to prove that I_fI_g is narrowly equivalent to I_h.

Without loss of generality, one can assume that f and g are united, since there will always be united forms in their classes. Let $f(x,y)=ax^2+bxy+ta'y^2$, $g(x,y)=a'x^2+bxy+tay^2$, $h(x,y)=aa'x^2+bxy+ty^2$, and let μ and μ' in $\mathbb{Q}\left(\sqrt{D}\right)$ whose norm has the same sign as a and a', respectively. We then have that the norm of $\mu\mu'$ has the same sign as aa' and that

$$I_fI_g=\left\langle a,\frac{-b+\sqrt{D}}{2}\right\rangle\mu\cdot\left\langle a',\frac{-b+\sqrt{D}}{2}\right\rangle\mu'=$$

$$=\left\langle aa',a\frac{-b+\sqrt{D}}{2},a'\frac{-b+\sqrt{D}}{2},\frac{b^2+(b^2-4taa')-2b\sqrt{D}}{4}\right\rangle\mu\mu'=$$

$$=\left\langle aa',a\frac{-b+\sqrt{D}}{2},a'\frac{-b+\sqrt{D}}{2},-b\frac{-b+\sqrt{D}}{2}-taa'\right\rangle\mu\mu'=$$

$$=\left\langle aa',\frac{-b+\sqrt{D}}{2},-b\frac{-b+\sqrt{D}}{2}\right\rangle\mu\mu'=\left\langle aa',\frac{-b+\sqrt{D}}{2}\right\rangle\mu\mu'=I_h,$$

where we used that, since F and G are united, $\gcd(a,a')=1$. □

Now we will see the relation between the units in orders and automorphisms of forms, but first, it would be useful to have another way to get a form from an ideal:

Definition 7.30. We say that $[\psi,\xi]$ is an *ordered basis* for a fractional ideal $\mathfrak{a}$ if $\mathfrak{a}=\langle\psi,\xi\rangle$ and ${}^{\sigma}\psi\xi-\psi\,{}^{\sigma}\xi$ is either positive or positive imaginary.

Observe that ${}^{\sigma}\psi\xi - \psi\,{}^{\sigma}\xi$ is always a pure imaginary number if the field is imaginary and it is always real if the field is real, so this definition only restricts the sign of the expression. If $[\psi, \xi]$ was not an ordered basis, $[\xi, \psi]$ will.

Proposition 7.31. *If $[\psi, \xi]$ is an ordered basis for a fractional ideal $\mathfrak{a}$, then $\mathfrak{a}$ is narrowly equivalent to I_f, being $f(x, y) = \frac{(\psi x+\xi y)({}^{\sigma}\psi x+{}^{\sigma}\xi y)}{N(\mathfrak{a})}$.*

Proof. Let $\omega = \frac{\xi}{\psi}$, let $ax^2 + bx + c$ $(a > 0)$ be its minimal polynomial and let $D = b^2 - 4ac$. On the one hand, from the definition of ordered basis, we have that $N(\psi)(\omega - {}^{\sigma}\omega)$ is either positive or positive imaginary, which implies that the sign of the radical in $\omega = \frac{-b\pm\sqrt{D}}{2a}$ is the same as $N(\psi)$. On the other hand,

$$f(x, y) = \frac{N(\psi)(x + \omega y)(x + {}^{\sigma}\omega y)}{\left|N\left(\frac{\psi}{a}\right)\right| N(\langle a, a\omega\rangle)} = \operatorname{sign}(N(\psi))a(x+\omega y)(x+{}^{\sigma}\omega y) = \operatorname{sign}(N(\psi))\left(ax^2 + bx + c\right).$$

Proposition 7.27 tells us that, depending on the sign of the radical and the norm, under the restriction concerning $N(\psi)(\omega - {}^{\sigma}\omega)$, $\mathfrak{a}$ is narrowly equivalent to $I_{ax^2+bxy+cy^2}$ or $I_{-ax^2-bxy-cy^2}$, which equals our I_f in either case. □

Proposition 7.32. *Let $\langle\psi, \xi\rangle = \langle\psi', \xi'\rangle$ and let $[\psi, \xi]$ be an ordered base. Suppose*

$$\begin{cases} \psi' = \alpha\psi + \beta\xi \\ \xi' = \gamma\psi + \delta\xi \end{cases}$$

is satisfied for some integers α, β, γ and δ. Then, $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ if and only if $[\psi', \xi']$ is also an ordered basis.

Proof. Since we are talking of a change of basis of $\mathbb{Z}$-modules, we automatically have that $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$. On the other hand,

$${}^{\sigma}\psi'\xi' - \psi'\,{}^{\sigma}\xi' = (\alpha\,{}^{\sigma}\psi + \beta\,{}^{\sigma}\xi)(\gamma\psi + \delta\xi) - (\alpha\psi + \beta\xi)(\gamma\,{}^{\sigma}\psi + \delta\,{}^{\sigma}\xi) = ({}^{\sigma}\psi\xi - \psi\,{}^{\sigma}\xi)(\alpha\delta - \beta\gamma)$$

implies the result. □

Proposition 7.33. *If $[\psi, \xi]$ and $[\psi', \xi']$ are ordered bases for the same fractional ideal $\mathfrak{a}$, and f and f' are the forms associated to them via Proposition 7.31, then f is transformed into f' by the action of the associated change of basis matrix $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$.*

Proof.

$$f'(x, y) = \frac{(\psi' x + \xi' y)({}^{\sigma}\psi' x + {}^{\sigma}\xi' y)}{N(\mathfrak{a})} =$$

$$= \frac{((\alpha\psi + \beta\xi)x + (\gamma\psi + \delta\xi)y)((\alpha\,{}^{\sigma}\psi + \beta\,{}^{\sigma}\xi)x + (\gamma\,{}^{\sigma}\psi + \delta\,{}^{\sigma}\xi)y)}{N(\mathfrak{a})} =$$

$$= \frac{((\alpha x + \beta y)\psi + (\gamma x + \delta y)\xi)((\alpha x + \beta y)\,{}^{\sigma}\psi + (\gamma x + \delta y)\,{}^{\sigma}\xi)}{N(\mathfrak{a})} = f(\alpha x + \beta y, \gamma x + \delta y)$$

□

Now, if ε is a unit of positive norm in $\mathcal{O}$ and $[\psi, \xi]$ is an ordered basis for $\mathfrak{a}$, then one can check that $[\varepsilon\psi, \varepsilon\xi]$ is another ordered basis for $\mathfrak{a}$, which leads us to the following theorem:

Theorem 7.34. *The group $\left(\mathcal{O}_D^+\right)^\times$ of units in $\mathcal{O}_D$ of positive norm is isomorphic to the group of proper automorphisms of any primitive form f of discriminant D.*

Proof. Let f be any primitive form of discriminant D, and let $\mathfrak{a}$ be a proper fractional ideal of $\mathcal{O}_D$, belonging to the narrow class corresponding to the proper class of f. We want to see that the group of automorphisms of f is isomorphic to the group $\left(\mathcal{O}_D^+\right)^\times$. Let $[\psi, \xi]$ be an ordered basis for $\mathfrak{a}$ and suppose without loss of generality that $f(x, y) = \frac{(\psi x + \xi y)({}^{\sigma}\psi x + {}^{\sigma}\xi y)}{N(\mathfrak{a})}$. This can be done since, if not equal, those two forms are properly equivalent, and their groups of automorphisms are conjugate (and so isomorphic).

To each unit ε of positive norm, which necessarily is $N(\varepsilon) = \varepsilon\,{}^{\sigma}\varepsilon = 1$, we can obtain a matrix in $\mathrm{SL}_2(\mathbb{Z})$ from the ordered bases $[\psi, \xi]$ and $[\varepsilon\psi, \varepsilon\xi]$ of $\mathfrak{a}$. Moreover, since

$$\frac{(\varepsilon\psi x + \varepsilon\xi y)({}^{\sigma}\varepsilon\,{}^{\sigma}\psi x + {}^{\sigma}\varepsilon\,{}^{\sigma}\xi y)}{N(\mathfrak{a})} = \frac{(\psi x + \xi y)({}^{\sigma}\psi x + {}^{\sigma}\xi y)}{N(\mathfrak{a})},$$

by the last proposition the matrix in question is a proper automorphism.

Conversely, given a proper automorphism $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ of f, we can consider $\psi' = \alpha\psi + \beta\xi$ and $\xi' = \gamma\psi + \delta\xi$, so that $[\psi', \xi']$ is also an ordered basis. Let $\omega = \frac{\xi}{\psi}$ and $\omega' = \frac{\xi'}{\psi'}$. By looking at the proof of Proposition 7.27, we see that both ω and ω' are roots of $f(x, 1)$, so we have $\omega' = \omega$ or $\omega' = {}^{\sigma}\omega$. Now, putting $x = 1, y = 0$ in the following equality

$$\frac{(\psi' x + \xi' y)({}^{\sigma}\psi' x + {}^{\sigma}\xi' y)}{N(\mathfrak{a})} = f(x, y) = \frac{(\psi x + \xi y)({}^{\sigma}\psi x + {}^{\sigma}\xi y)}{N(\mathfrak{a})},$$

we see that $N(\psi) = N(\psi')$, which implies that $\frac{\xi'}{\psi'} = \omega' = \omega = \frac{\xi}{\psi}$, since the definition of ordered basis ensures both $N(\psi)(\omega - {}^{\sigma}\omega)$ and $N(\psi')(\omega' - {}^{\sigma}\omega')$ are positive or positive imaginary. That means that there exists some ε, in principle in $\mathbb{Q}\left(\sqrt{D}\right)$, such that $\varepsilon = \frac{\psi'}{\psi} = \frac{\xi'}{\xi}$ $(= \alpha + \beta\omega)$. However, since $\varepsilon\,\langle\psi, \xi\rangle = \langle\psi', \xi'\rangle \subseteq \langle\psi, \xi\rangle$, we have that $\varepsilon \in \mathcal{O}_D$ because of the properness of $\mathfrak{a}$. Finally, we have that $N(\varepsilon) = 1$ since $N(\psi') = N(\psi)$. Therefore, $\varepsilon \in \left(\mathcal{O}_D^+\right)^\times$.

Clearly, those two maps are mutually inverse. It is easy to see that multiplication of units translates into nestings of linear equations, which turn into matrix multiplication, so the maps

are isomorphisms indeed. $\square$

Recall that we had another bijection between proper automorphisms and solutions to Pell's equation $x^2 - Dy^2 = 4$, and we even established a group law on this last set, which resembled some bizarre multiplication. This is no accident:

First, consider a proper automorphism of $f(x,y) = ax^2 + bxy + cy^2$ ruled by a solution to Pell's equation, $U(X,Y) = \begin{pmatrix} \frac{X-Yb}{2} & -cY \\ aY & \frac{X+Yb}{2} \end{pmatrix}$. This automorphism induces a unit in $\mathcal{O}_D$, through an ordered basis of $I_f = \left\langle a, \frac{-b+\sqrt{D}}{2} \right\rangle \mu$. One natural choice would be $\left[\mu a, \mu\frac{-b+\sqrt{D}}{2}\right]$, which is proper since $N(\mu)\left(a\frac{-b+\sqrt{D}}{2} - a\frac{-b-\sqrt{D}}{2}\right) = N(\mu)a\sqrt{D}$ is either positive or positive imaginary. However, it is more convenient to use the basis $\left[\mu' c, \mu'\frac{b+\sqrt{D}}{2}\right]$ of $I_g = \left\langle c, \frac{b+\sqrt{D}}{2} \right\rangle \mu'$, being $g(x,y) = cx^2 - bxy + ay^2$ properly equivalent to $f(x,y)$. Choosing this basis, we obtain $\varepsilon = \frac{X-Yb}{2} + (-cY)\frac{b+\sqrt{D}}{2c} = \frac{X-Y\sqrt{D}}{2}$. Moreover, the group law $(X,Y) \circ (X',Y') = \left(\frac{XX'-DYY'}{2}, \frac{XX'+YY'}{2}\right)$ on the set of solutions becomes more meaningful:

$$\frac{X - Y\sqrt{D}}{2} \cdot \frac{X' - Y'\sqrt{D}}{2} = \frac{\left(\frac{XX'-DYY'}{2}\right) - \left(\frac{XX'+YY'}{2}\right)\sqrt{D}}{2}$$

8 Conclusions and further investigations

This degree thesis ends here. Following the works of great mathematicians, we have studied integer binary quadratic forms. We saw how the notion of equivalence arises when considering representation of a number by different forms, and we also defined what it means to be 'proper'. We managed to identify the proper classes using some special kind of forms, called reduced, half-reduced or diminished, depending on the discriminant. This helped us to identify whether two forms are properly/improperly equivalent, and to be able to compute the class number $h(D)$ for any given discriminant D.

After that, we also defined the notion of composition of forms, and we saw that this operation gave a group structure to the set of proper classes. For square discriminants, we determined the structure of the class group exactly: for $D = d^2 > 0$, it is isomorphic to $(\mathbb{Z}/d\mathbb{Z})^\times$; for $D = 0$, it is trivial. For non-square discriminants, we found another group isomorphic to it, the narrow ideal class group.

Apart from that, we also established some connections of quadratic forms with other branches of mathematics. We saw how the group of automorphisms turns out to be isomorphic to the group of units in a quadratic field, and is also in bijection with the units of $x^2 - Dy^2 = 4$. We saw how the tail of continued fraction of $\frac{-b+\sqrt{D}}{2}a$, which necessarily repeated, was displaying what were the transformations between neighbouring half-reduced forms used by the algorithm. We even saw a proof (definitely not the simplest, though) of the fact that the group $\mathrm{SL}_2(\mathbb{Z})$ is generated by two elements: $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

However, there has been much more progress on the topic that the one presented here. For instance, since ideals of an order are also lattices in $\mathbb{C}$, we can consider the elliptic curve associated to that lattice and how they behave under the equivalence defined on the lattices. This connection with elliptic curves, and so with L-series, ultimately allows to compute $h(D)$ in terms of an L-series, [5, p. 188][4, p. 83] or [3, p. 215]. Using this, one can prove that there are only finitely many negative discriminants with class number 1, which we presented in Example 2.6.

In addition to all the progress on integer forms, nowadays, as we said in the introduction, the study is not limited to integer coefficients, further progress concerning quadratic forms in other rings/fields have been made. Also, the study of classes of ideals has been extended to other number fields, not just quadratic.

Finally, I wish that this degree thesis has been a decent introduction to the world of integer quadratic forms, and I want to thank my tutor, Prof. Jordi Quer, for his patience with me and also my parents and friends for their support.

References

[1] C. F. Gauß. *Disquisitiones arithmeticae*, Springer-Verlag, New York, 1986. Translated and with a preface by Arthur A. Clarke, Revised by William C. Waterhouse, Cornelius Greither and A. W. Grootendorst and with a preface by Waterhouse.

[2] David A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, John Wiley & Sons, Inc., New York, 1989.

[3] Henri Cohen, *A course in computational algebraic number theory*, Graduate texts in mathematics, vol. 138, Springer-Verlag, Berlin, 1993.

[4] D.A. Buell, *Binary Quadratic Forms, Classical Theory and Modern Computations*, Springer-Verlag, New York, 1989.

[5] J. Buchmann and U. Vollmer, *Binary quadratic forms: An algorithmic approach*, Algorithms and Computation in Mathematics, vol. 20, Springer-Verlag, Berlin, 2007